SUPPORTING FLIGHT CONTROL FOR UAV-ASSISTED WILDERNESS SEARCH AND RESCUE THROUGH HUMAN CENTERED INTERFACE DESIGN

by

Joseph L. Cooper

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computer Science

Brigham Young University

December 2007

Report Documentation Page					Form Approved OMB No. 0704-0188
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE **DEC 2007**		2. REPORT TYPE			3. DATES COVERED **00-00-2007 to 00-00-2007**
4. TITLE AND SUBTITLE **Supporting Flight Control for UAV-Assisted Wilderness Search and Rescue Through Human Centered Interface Design**					5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)					5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Brigham Young University ,Department of Computer Science,Rexburg,ID,83460**					8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT **Inexpensive, rapidly deployable, camera-equipped Unmanned Aerial Vehicle (UAV) systems can potentially assist with a huge number of tasks. However, in many cases such as wilderness search and rescue (WiSAR), the potential users of the system may not be trained as pilots. Simple interface concepts can be used to build an interaction layer that allows an individual with minimal operator training to use the system to facilitate a search or inspection task. We describe an analysis of WiSAR as currently accomplished and show how a UAV system might fit into the existing structure. We then discuss preliminary system design efforts for making UAVenabled search possible and practical. Finally, we present both a carefully controlled experiment and partially structured field trials that illustrate principles for making UAV-assisted search a reality. Our experiments show that the traditional method for controlling a camera-enabled UAV is significantly more difficult than integrated methods. Success and troubles during field trials illustrate several desiderata and information needs for a UAV search system.**					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**129**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Joseph L. Cooper

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Michael A. Goodrich, Chair
Date	Bryan S. Morse
Date	Phillip J. Windley

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Joseph L. Cooper in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date	Michael A. Goodrich Chair, Graduate Committee
Accepted for the Department	Parris K. Egbert Graduate Coordinator
Accepted for the College	Thomas W. Sederberg Associate Dean, College of Physical and Mathematical Sciences

ABSTRACT

SUPPORTING FLIGHT CONTROL FOR UAV-ASSISTED WILDERNESS SEARCH AND RESCUE THROUGH HUMAN CENTERED INTERFACE DESIGN

Joseph L. Cooper

Department of Computer Science

Master of Science

Inexpensive, rapidly deployable, camera-equipped Unmanned Aerial Vehicle (UAV) systems can potentially assist with a huge number of tasks. However, in many cases such as wilderness search and rescue (WiSAR), the potential users of the system may not be trained as pilots. Simple interface concepts can be used to build an interaction layer that allows an individual with minimal operator training to use the system to facilitate a search or inspection task. We describe an analysis of WiSAR as currently accomplished and show how a UAV system might fit into the existing structure. We then discuss preliminary system design efforts for making UAV-enabled search possible and practical. Finally, we present both a carefully controlled experiment and partially structured field trials that illustrate principles for making UAV-assisted search a reality. Our experiments show that the traditional method for controlling a camera-enabled UAV is significantly more difficult than integrated

methods. Success and troubles during field trials illustrate several desiderata and information needs for a UAV search system.

ACKNOWLEDGMENTS

My wonderful, loving wife, Amanda, has sacrificed, encouraged, and patiently helped me all the way through the graduate program. My daughter, Emily, gave me inspiration. My parents and siblings gave their support. Dr. Michael Goodrich, my advisor, carefully helped me through the creative discovery process and ensured that this thesis could become a reality. Colleagues in the Human Centered Machine Intelligence Lab made it fun.

Much of this work represents collaborative effort with the MAGICC lab and Dr. Bryan Morse and his students at BYU as well as Dr. Julie Adams and Curtis Humphrey at Vanderbilt University. Most of all, my Father in Heaven has blessed and supported me all my life and in all aspects of my life; my thesis research is no exception. This research has been funded in part by NSF grant number IIS-0534736 and in part by ARL-HRED contract number DAAD19-01-C-0065. I am very grateful for all of the support and assistance received from so many different directions.

Contents

1 Introduction **1**

1.1 Background on UAVs . . . 1

1.2 Inexpensive air support for wilderness search and rescue . . . 3

1.3 Thesis statement . . . 4

1.4 Overview . . . 4

2 Related work **7**

2.1 Current Flight Systems . . . 7

2.2 General Human Factors . . . 13

2.3 Ecological design . . . 15

2.4 Feature focused research . . . 16

2.4.1 Perspective . . . 17

2.4.2 Attention . . . 19

2.4.3 Organization and Layout . . . 20

2.5 Task analysis and interface design . . . 20

3 Task Analysis **23**

3.1 Goal Directed Task Analysis . . . 23

3.1.1 Stage Preparation - Goal 1.0 . . . 26

3.1.2 Missing Person Description - Goal 2.0 . . . 26

3.1.3 Search Plan - Goal 3.0 . . . 28

3.1.4 Execution of Search Plan - Goal 4.0 . . . 33

3.1.5 Recovery - Goal 5.0 and Debriefing - Goal 6.0 36
3.2 Information Flow 36
3.3 Activity Analysis and Task Breakdown 39
3.3.1 UAV-Enabled WiSAR: Task Breakdown 39
3.3.2 Deployment, Monitoring, and Retrieval 40
3.3.3 Searching for Evidence 43
3.3.4 Constrain Search 45
3.3.5 Integration into Existing Process 48

4 System Design 51
4.1 Platform 51
4.1.1 Airframe 51
4.1.2 Ground system hardware 53
4.2 Automation and Abstraction 54
4.2.1 Autopilot 55
4.2.2 Ground station automation 56
4.3 Information presentation 58
4.3.1 Intuitive presentation 59
4.3.2 Perspective 64
4.3.3 Attention 70
4.4 Interface evolution 72
4.5 Software architecture 76

5 Validation 81
5.1 Small-scale experiments 81
5.1.1 Multi-window target detection study 82
5.1.2 Path recall study 85
5.2 Perspective experiment 87

5.2.1 Design . . . 88
5.2.2 Results . . . 92
5.3 Field trials . . . 99
5.3.1 Neglect tolerance . . . 101
5.3.2 Persistent, enhanced, terrain referenced imagery . . . 103
5.3.3 Reliable communication lines . . . 104
5.3.4 Organized plan . . . 105

6 Conclusion and Future Work 107

6.1 Conclusion . . . 107
6.2 Future work . . . 109
6.2.1 Multi-agent interface extension . . . 109
6.2.2 Integration with mosaic . . . 110
6.2.3 Full 3D interaction model . . . 110
6.2.4 Playback functionality . . . 111
6.2.5 Sophisticated 3D path planning . . . 111
6.2.6 Airspace integration / Meeting FAA regulations . . . 112
6.2.7 Integration with other WiSAR technology . . . 112

Chapter 1

Introduction

This thesis presents research toward using camera-equipped Unmanned Aerial Vehicles (UAVs) to support Wilderness Search and Rescue (WiSAR) efforts. Accomplishing this goal not only has the potential to do a great deal of good, but also brings up many interesting problems.

1.1 Background on UAVs

UAVs have been used for various military tasks since the time of World War I [57]. Imagery capability was first introduced to remotely piloted aircraft in the 1950s when the Ryan Aeronautical Company adapted radio-controlled drones used for target practice to carry a camera and fly a preprogrammed course [60]. Ryan Aeronautical hoped to develop a technology that would provide intelligence imagery of Soviet installations without endangering a human pilot. Recently, military operations have come to rely heavily on UAVs. The Hunter, Shadow, and Predator drones provide invaluable intelligence and even munitions deployment for military activities such as operation Iraqi Freedom and operation Enduring Freedom. Researchers are now recognizing that many of the advantages camera-equipped UAVs provide for military service may also extend to a number of civilian purposes from border patrol and meteorology to bridge inspection and journalism [26]. WiSAR is one particular area in which camera-equipped UAVs may continue to serve society.

However, UAV technology is not trivially introduced as a solution to a problem. Just as with manned aircraft, a UAV system must overcome the complications associated with flight, balancing weight and aeronautical design with functionality. Sophisticated system design can provide advanced capability, but may also introduce complications and potential human error. System design must provide the proper set of abilities to enable an operator to accomplish the task and then expose the abilities through the system interface such that accomplishing the task is feasible within human limitations.

Because UAVs are remotely operated, many of the cues that pilots traditionally rely on are not present. The operator is prone to lose track of where the craft is and what it is doing. The separation of the operator from the craft makes it critical for a UAV system to appropriately present necessary information to the operator. Some early UAV systems relied almost exclusively on the video signal for communicating the state of the craft, an approach that has been equated with navigating through a soda straw [68]. It is quite difficult to get a feeling for scale and robotic footprint exclusively through video [22]. It may be even more difficult for the operator to anticipate the *future* state of the craft. Understanding the current state of the craft, recognizing its with relationship with the world, and predicting the future consequences of operator decisions are often combined into a general concept known as *situation awareness* [16].

Situation awareness is critical for all stages of flight although the precise knowledge requirements for different tasks differ. The problem of maintaining situation awareness is exacerbated by the fact that for a search task, the operator's attention is partially devoted either to inspecting the imagery or to interacting with someone else (such as a sensor operator in charge of monitoring the video) in order to refine the imagery. Some of the operator burden can be relieved through automation of the UAV, but this also adds an additional system for the operator to understand and anticipate and may cause difficulties by disconnecting the operator so far from

the task that when a critical decision must be made, the operator has insufficient understanding to make an appropriate choice [4].

1.2 Inexpensive air support for wilderness search and rescue

WiSAR is a demanding field of work. That it can also be rewarding work is evidenced by the fact that the Utah County Sheriff's Search and Rescue Team is composed almost entirely of volunteers who are expected to expend thousands of dollars of personal resources for rescue equipment and be on call 24 hours a day, 365 days a year [12]. WiSAR volunteers (also referred to as "first responders" in this document) may be called to perform their duties in mountains, deserts, lakes, and other terrain that requires special equipment to cover in a timely manner. Team members occasionally expose themselves to risks inherent in negotiating hazardous environments in the course of duty.

Private, manned aircraft are occasionally used to assist with a search, but even small manned aircraft may take a relatively long time to get into the air and are then limited by minimum altitude and airspeed constraints for the safety of the pilot and others. Manned aircraft may also be prohibitively expensive to run. An inexpensive, easily portable alternative is needed to provide aerial imagery to assist in the search effort. Small, camera-equipped UAVs have the potential to provide an affordable alternative that can be carried in-hand to the search area and flown inexpensively to quickly cover a site visually without disturbing other signs such as scent trails used by canine tracking teams. In Chapter 3 we review details of WiSAR much more thoroughly and discuss how camera-equipped UAVs may be used to facilitate the process.

1.3 Thesis statement

By appropriately combining robot autonomy and interface design to support situation awareness, we can create a UAV control interface that non-pilot operators can use to successfully execute an aerial search task after minimal instruction. The interface provides support for major subtasks of a WiSAR operation through combinations of autonomy and various methods of information presentation.

1.4 Overview

In addressing the issue of designing a UAV system capable of supporting WiSAR, we begin with a review of relevant literature. This includes other flight systems as well as similar research for remotely operated ground vehicles. We also review interface design issues and human subject studies similar to those reported in Chapter 5.

We use formal task analysis to capture WiSAR as it is currently accomplished. This analysis focuses specifically on goals, information requirements for those goals, and a model of information flow in WiSAR. The analysis results inform a discussion on the potential for introducing UAV technology into WiSAR along with issues to be addressed in order to make it possible and productive.

Such issues include appropriate interface and automation for using a UAV to meet the information requirements for major search and rescue goals. We discuss the design and implementation of an interface intended to meet the constraints imposed by UAV-enabled WiSAR. Controlling a UAV from a single-display ground station can be difficult and requires careful design for adequate information presentation. Because it was a significant part of this project, our discussion on interface design includes a brief discussion of software architecture that allows the interface to accomplish both control in the field and experimental testing in the lab. Some design decisions for the system are justified based on prior or related work. Other decisions are validated

through experimental or empirical testing. Still other system features remain untested and must be addressed in future work.

The experimental and empirical validation we have performed is noteworthy. Several simple, preliminary experiments show some basic limitations and strengths of human cognition and abilities. A more thorough study performed in simulation using several different virtual perspectives for a search task illustrates the strength of an ecological design and highlights principles for information presentation in a WiSAR UAV interface.

Several field trials performed during this research give the simulated experiments and system design a grounding in reality. Experiences in the field expose difficult problems as well as promising directions for future work. We conclude with a discussion of research that other researchers are currently pursuing as well as some problems that still remain untouched.

Chapter 2

Related work

To work toward building a supportive UAV system for WiSAR, this thesis builds on research from many areas and disciplines. After briefly discussing modern flight control systems, both manned and unmanned, we will review some general principles of human factors applied to human-robot interaction. We will then examine human-robot interfaces designed to support the challenges of remote operation. A significant amount of interface research focuses on specific interface features and principles—so much that we can only cover a small subset of relevant studies. Specifically, we will discuss perspective in ecological design and principles of attention and organization. Finally, we will review the use of task analysis to inform system design.

2.1 Current Flight Systems

When UAVs first began to be used, they were essentially missiles with a little bit of control. Perhaps the first UAV interface that provided inflight information and control was in the 1950s. Operators used a grease pencil to trace the path of the UAV on a radar screen and used a simple radio connection to make basic flight adjustments [60]. As UAVs became less like missiles and more like planes, it was natural to adopt control ideas from manned flight. The typical modern ground-control system is designed to imitate, at least partially, a traditional manned aircraft control paradigm.

Figure 2.1: Boeing 737 captain's instruments
(from http://www.b737.org.uk)

In this digital age, the field of flight control, in general, is still based largely on analog devices. A pilot controlling an aircraft through direct manipulation of control surfaces requires certain information to be successful [61]. Even in a modern, computer-equipped cockpit, information on the screen is often presented using digital representations of analog dials and gauges that were originally connected directly to mechanical devices. These dials and gauges are comfortable and familiar to trained pilots, but may be foreign and confusing for the uninitiated. Figures 2.1 through 2.4 show components from a typical commercial aircraft cockpit with gauges, lights, and switches for controlling and monitoring the many sub-systems on a large aircraft. Smaller aircraft have fewer systems, but still have a similar base set of components [61].

Perhaps the most prominent example of a UAV control system modeled after manned aircraft is the United States Air Force Predator UAV, currently flown in military reconnaissance and munitions deployment. Despite the differences that arise through remote operation and computer-assisted flight, the Predator ground control station is designed to closely replicate an aircraft cockpit in many respects (Figure 2.6) and is operated exclusively by qualified air force pilots [9].

Figure 2.2: Boeing 737 center panel (from http://www.b737.org.uk)

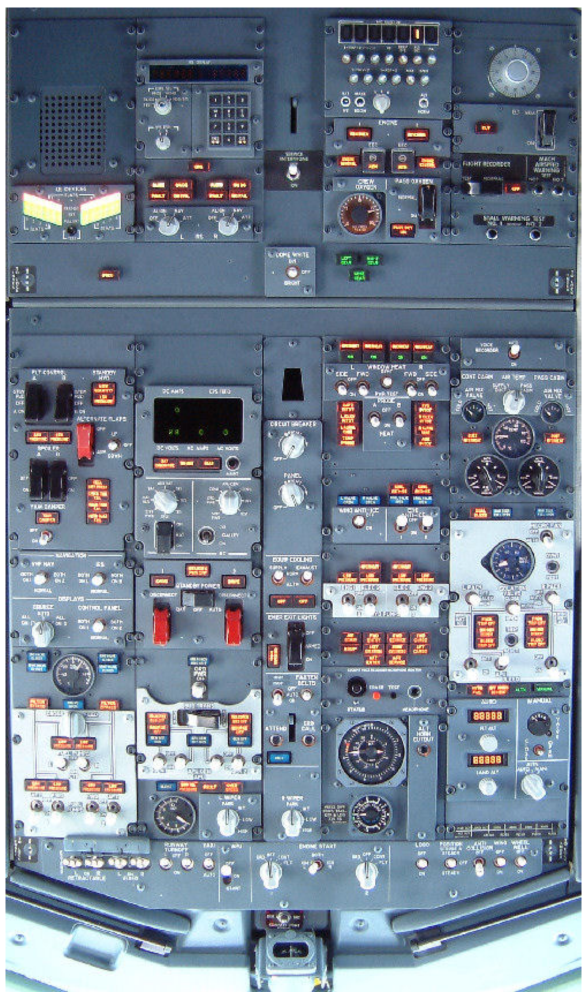


Figure 2.3: Boeing 737 overhead panel (from http://www.b737.org.uk)

Figure 2.4: Boeing business jet glass-cockpit
(from http://www.b737.org.uk)

Figure 2.5: External view of the Predator ground station
(Photo by Nathan Rackliffe)

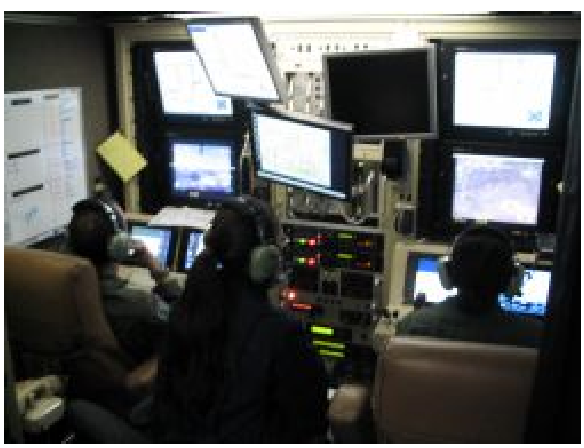


Figure 2.6: Predator display and control
(Photo by Nathan Rackliffe)

Although controls may be placed in the same configuration as in a manned aircraft, pilots find that many tasks are more difficult because of the lack of peripheral, audio, and vestibular cues (often referred to as “flying by the seat of your pants”) [68]. The task of piloting a manned aircraft is not the same as remotely controlling a UAV and the appropriate control model for one may not translate well to the other. The Air Force reports a disproportionately high level of accidents with UAVs. The reports frequently blame the pilot for the accident, but the design of the control system is at least partially at fault [66]. For example, in one case, a pilot used a three-key sequence that typically executes a very common flight procedure. However, because the interface was in an unexpected state, the key sequence instructed the craft to deprogram itself mid-flight. Lobotomized, the craft stopped all communication with the ground-station and crashed [10]. Although it is true, as the report claimed, that the pilot did not follow procedure of always verifying the interface mode before issuing a command, the interface should make mode more obvious so that confusion is less common [48] and the flight control interface should not expose commands that are never supposed to be used while the craft is in flight. Many other UAV interface systems exist that are less extreme than the Predator system but are similar to each other in their attempt to incorporate manned flight controls into a ground-based computer display (Figures 2.7 through 2.10). Ruck referred to the typical UAV interface as a system designed by a 23 year old engineer just out of college in a way that makes sense to himself but to no one else [46].

Although many systems exist for controlling UAVs, nothing seems to exist that meets the limited pilot training, high-mobility, and low-cost constraints of the WiSAR domain. The WiSAR volunteer may not have extensive flight training and so a cockpit inspired interface may be overly complex. Furthermore, for mobility reasons, the trailer-load of equipment (Figure 2.5) necessary to duplicate a cockpit is not practical for WiSAR. Even those UAV interface systems designed to run on a

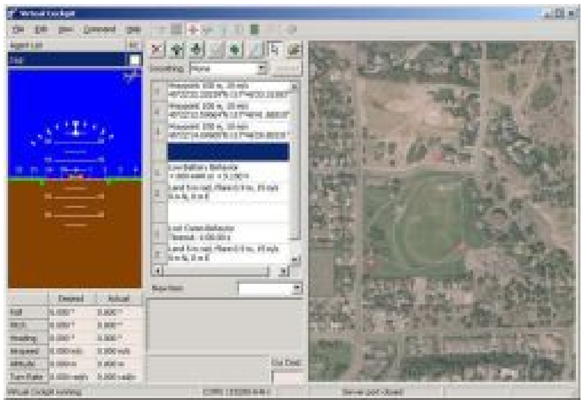


Figure 2.7: BYU Magicc lab “Virtual Cockpit” interface

Figure 2.8: Applied Research Associates TACMAV interface
(from http://www.ara.com/mpsp/ECD/seg/TACMAVOverview.htm)

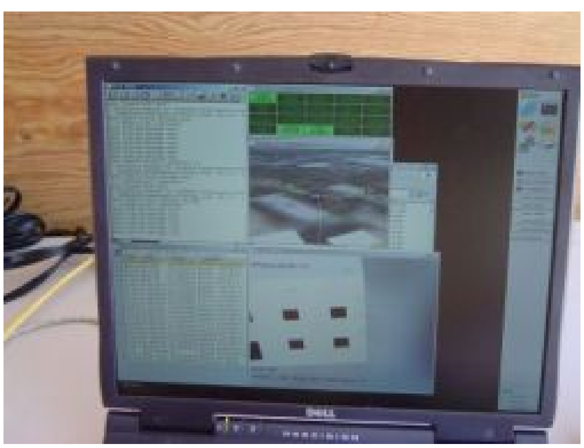


Figure 2.9: Georgia Tech GCS
(from http://uav.ae.gatech.edu/)

Figure 2.10: University of Bologna GCS
(from http://www.ingfo.unibo.it/)

single computer with a basic point and click interface still often have a steep learning curve and induce a heavy cognitive load.

2.2 General Human Factors

Human operators have strengths, weaknesses, and general tendencies that must be accounted for when designing a human-robot interface. Human factors research documents several phenomena and requirements that are relevant for our work. Perhaps the most important humans factors principle is encapsulated in the saying, "To err is human...". In spite of training and talent, people can still become tired, distracted, and confused as in the case of the Predator accident described in Section 2.1. Although this common sense statement seems obvious, system designers and engineers are prone to forget that the operator can not be expected to perform perfectly, constantly, and consistently.

One common source of error in remotely operating a robotic system is a lack of understanding of the system state. How well the operator understands the past, present, and projected behavior of the system is commonly known as *situation awareness* [16]. If the operator has an incorrect understanding of the current system state,

he or she is much more likely to make poor decisions that negatively affect performance.

It is generally agreed that situation awareness is very important, but can be difficult to define for a specific case and even more difficult to measure. Several methods have been proposed for quantifying situation awareness during a task [55], but others argue that the measurement process is flawed because the measurement techniques influence the actual awareness through interruption and prompting. It can be argued that the only measure of awareness that is really important is performance. If a subject can consistently achieve high performance and avoid catastrophic failure with a particular system under a wide range of operating conditions, the rest does not matter. We assume in this thesis that higher performance implies more informed decision making and better awareness.

Related to the principle of situation awareness are the ideas of *mode confusion* and *change blindness*. Mode confusion [48] occurs when the system is not in the state that the operator expects. Confusion about the system's operating mode can lead the operator to misinterpret information presented by the interface or issue one command when intending another. These misinterpretations can lead to catastrophic errors. Mode confusion can occur if the two modes appear similar and the operator forgets which mode he or she last used. It may also occur if the system can autonomously change modes and the operator does not notice. Change blindness can make this more common that one might expect. Change blindness is an interesting phenomenon where large changes can occur and if an individual is not attending to the particular thing that changed, he or she may not notice [52].

Requiring a system operator to constantly attend to an interface or anything else is impractical because of the principles of *cognitive work* and *neglect tolerance*. Even a task as simple as monitoring video from a security camera for any length of time can be fatiguing and performance inevitably declines [4]. Attending, mentally

transforming, and processing data require quantifiable effort and there are limits to what one can accomplish in a given length of time. This leads to the need for neglect tolerance. Often used to describe how many robots a single operator can successfully operate, neglect tolerance measures how long, on average, a single robot system may be neglected before performance degrades below some critical point [20]. Although our current intent is only to provide control for a single UAV, WiSAR volunteers can be expected to experience many distractions. Furthermore, the video system and the flight system are sufficiently separated and cognitively demanding that most UAV interfaces assign them to separate operators. With limited manpower, two UAV operators may not be an option for WiSAR. It is therefore useful to be aware of the neglect tolerance of both systems to know how well a single operator can expect to use both while filling other responsibilities accessory to the robotic control task.

2.3 Ecological design

Applying general human factors knowledge to interface design has led many groups to employ ecological design for improved situation awareness. The principle of ecological design is to integrate sensor information, video, and other previously acquired information into a single natural interface. This idea and its variations go by many names: virtual, mixed, or augmented reality, virtual or synthetic environments, and augmented virtuality. For a more complete discussion of ecological design and the finer distinctions between its different labels, see [35]. The point is to improve situation awareness and reduce cognitive workload by communicating the situation in a graphical manner more easily understood than a collection of dials, lights, and numeric displays.

Considerable evidence shows that ecological design can be beneficial for remotely operating robotic systems. Ricks found that it is easier to control a remote ground vehicle with an ecological interface than with a conventional interface using

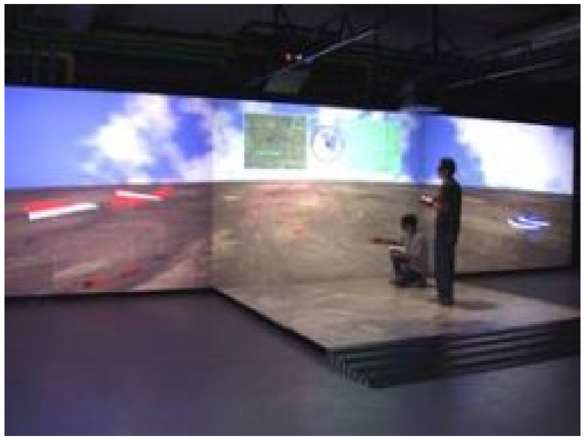


Figure 2.11: Iowa State Virtual Reality UAV Interface
(from http://www.vrac.iastate.edu/~sannier/VirtualTeleop/)

separate displays for separate sensors [45]. A handful of groups are also working on ecological UAV interfaces. The VRAC group at Iowa State University has developed and experimented with a virtual reality, immersive interface shown in Figure 2.11. With this interface, Knutzon conducted quantitative and qualitative user-studies and found that situation awareness was positively correlated with the increased field-of-view provided by the synthetic environment [31]. Drury et al. also found that displaying the video from a UAV in context using a synthetic environment improved perception of the video over raw video [14].

It must be noted, however, that Smallman and St. John have found that increased realism typically makes a more impressive looking interface, but not always a more effective interface [53]. Some display techniques, while visually appealing, tend to obscure information rather than make it available.

2.4 Feature focused research

Using an ecological model is one of the many design decisions to be made in developing a system for UAV-assisted WiSAR. A tremendous amount of human-computer interaction research explores the effects of various specific features in an interface.

This work reveals specific principles that we employ to accomplish our design goals. From such a large body of research, we can only discuss a relatively small sample of the relevant studies. In this section we explore literature related to presenting the synthetic environment, controlling the flow of information to the operator, and organizing the interface for usability.

2.4.1 Perspective

With an ecological design, the synthetic environment model is responsible for communicating a significant amount of information about the terrain, the craft, the relationship between them, and other spatial information. Rendering three-dimensional information to a computer display requires a "virtual camera" that defines how to accomplish the projection from 3D to 2D. The virtual camera combines frame of reference, perspective, and field of view to generate a 2D image of the scene (see [7]). The virtual camera controls how the synthetic environment is displayed to the operator and consequently what information is available and what information is obscured. For example, if the virtual camera is looking down at the synthetic terrain, variations in terrain altitude are less visible, but horizontal distances are easier to see.

Many studies claim to compare 2D interfaces against 3D interfaces for accomplishing some flight task (e.g., [3, 30, 64]). Stating the problem this way fails to capture the fact that all interfaces displayed on a computer screen are 2D. Any portrayal of the craft and/or terrain must be a two-dimensional projection of a three-dimensional space. The distinction is strictly one of axis alignment. One such study stated that the only way to make a "fair comparison" between 2D and 3D was to give the 2D interface two different viewpoints (top-down and forward) [64]. What this study called a 3D viewpoint placed the virtual camera somewhere between directly above and directly behind the craft. The presentation with two viewpoints has more information available than any single viewpoint can; so it is not surprising that

the study found that the 3D interface performed worse than the 2D because of the ambiguity from the 3D projection. Every projection from three-space to two-space introduces ambiguity as information is compressed along one axis. The top-down perspective leaves altitude ambiguous. The forward perspective leaves depth ambiguous. A projection that is not aligned with a labeled axis will still introduce just as much ambiguity.

From the many studies done comparing usability of different display perspectives, the one general conclusion has been that task performance is, in fact, related to display, but the exact relationships are uncertain [3]. This seems to result from confounding differences in the way interfaces used for comparison are presented. Many other factors besides perspective play a major part in performance. The most reasonable and believable conclusion of all these studies is that the most important thing is for necessary information to be available and accessible in one way or another [54]. An operator needs certain information to accomplish a task well. Although a given perspective may make certain information ambiguous, other interface elements can compensate for that.

It may not be possible to develop "one true interface" that is ideal for every type of task the WiSAR volunteer may perform. However, there are interface presentation methods that are more or less appropriate for particular types of tasks and combinations of autonomy [50]. Wickens suggests that an immersed view (first person) is more effective for tasks involving local movement and a plan view (2D fixed-orientation map) is more effective for tasks involving understanding spatial relationships [64]. Because the WiSAR operator will need to perform both types of task, the UAV interface should include both perspectives.

If a single display interface has the ability to display multiple perspectives, it either needs multiple windows to show them simultaneously or it needs a way to transition between the different perspectives. Plumlee and Ware describe several methods

for manipulating a virtual camera in a synthetic environment [38, 39]. They find that smooth transitions can help maintain knowledge obtained from one perspective for use in another. In other words, smooth transitions between perspectives can help avoid operator disorientation.

A related but separate concept relates to the presentation of a video signal to the operator. With a camera-equipped UAV, the entire purpose of putting the craft into the air is to obtain imagery. When using a single operator interface design or if the flight path must change reactively as imager is acquired, it is particularly important to make the video information available to the operator. Plumlee and Ware explore methods for connecting a separate video window to a craft model in a synthetic environment and found that tethers (lines drawn between the craft and the corners of the video window) did not help much. What did help was rotating the world to maintain a track-up perspective and showing a "proxy" in the synthetic environment which indicates where the camera is pointing [37]. Drury et al. and Calhoun et al. both found that displaying video surrounded by some synthetic terrain improves understanding of the video [8, 14].

2.4.2 Attention

With many sources of information competing for the operator's attention, it is important to be aware of distractions and information accessibility in an interface. The problem of change blindness can also be partially mitigated by controlling information elements to attract attention, but these techniques must be used carefully.

Controlling saliency of interface elements leads to lower clutter and therefore less distraction, but keeps information available in case it is necessary [23]. The key is to have information available when it is needed. Ideally, only the needed information is available. However, since different operators use information differently, we must compromise.

Display cluttering occurs when an interface tries to present too much information at once, or the information is not structured such that the user can integrate it effectively into a mental model [62]. Cluttered displays hinder the operator from focusing where necessary and make it difficult to find, fuse, and use information. Information and controls can be buried in menus and dialogs to declutter the screen, but the risk is that critical information or control will not be available when it is needed [62].

2.4.3 Organization and Layout

Literature examining the effects of clutter have reached the somewhat obvious conclusion that increased clutter makes an interface more difficult to use. Likewise complicated menu structures with randomly grouped functions are more difficult than simple menus with functions organized according to function. Hiding or separating interface elements may also lead to increased delay and mental workload because it requires the operator to remember where information and controls are and how to find [17] and interpret them [62]. This can introduce hesitation and errors at critical moments. According to the Proximity Compatibility Principle [62], it is important to locate interface elements with similar function or feedback close together and those which are unrelated should be far apart. Another method for reducing clutter is to segregate information and control according to different modes and only provide those which are relevant to the current mode. However, this has the potential of introducing mode confusion [48].

2.5 Task analysis and interface design

In Chapter 3 we discuss a formal task analysis used to inform the design of our control interface. Saja emphasizes the importance of engendering a correct cognitive model of the system so that the operator understands what options are available and what

their consequences will be [47]. Formal analysis seems to be most frequently applied to tasks where certain failures can be catastrophic such as a nuclear plant [59]. Using formal analysis of a task to determine different task phases, changing information requirements, and information flow may not be a particularly new idea, but it must be adapted for the needs of each domain to which it is applied; see [13, 58].

Chapter 3

Task Analysis

A critical part of developing a system capable of assisting with WiSAR is ensuring that the system is designed to perform a function that is actually helpful to first responders. Formal analysis techniques provide a structured framework for systematically reviewing goals, information flow, resource allocation, and other information about accomplishing a task. A thorough analysis of the WiSAR domain allows us to see how the task is currently accomplished and how technology may fit into the current structure to fill a productive role. Furthermore, an analysis of information needs allows us to design the interface to appropriately support specific tasks. Together with Curtis Humphrey and Julie Adams of Vanderbilt University, we have studied the WiSAR domain using two task analysis techniques: Goal Directed Task Analysis (GDTA) [15] and Cognitive Work Analysis [13, 58]. In this thesis, we present the results from the GDTA together with conclusions from the full analysis and implications for UAV system design. Julie and Curtis contributed to the writing in this chapter as part of a collaborative technical report [1]. Portions of this chapter are also in [19] and [14].

3.1 Goal Directed Task Analysis

We performed the GDTA in order to understand the wilderness search process by identifying the WiSAR team goals, decisions, and ideal information requirements.

GDTA is not bound to the current system, and permits identification of potential system improvements. The GDTA has four stages: goal hierarchy development, conducting interviews, developing the goal-decision-SA (situation-awareness) structure, and obtaining feedback. Subject matter experts, Ron Zeeman, Kent Compton, and Brian Buss kindly provided information and reviewed the analysis results. All three have worked in the past or are currently working on the Utah County Search and Rescue team.

The GDTA identifies six unique high-level WiSAR goals along with a number of subgoals, decision questions, and information requirements. A graphical representation of the GDTA, developed together with Curtis Humphrey, is presented in Figure 3.1. The overall goal is the rescue or recovery of a missing person.

The first responders have three main priorities that they strive to achieve. The first priority is their own personal safety. Although this goal is emphasized in subgoal 4.3, it is a primary consideration for all stages of WiSAR. Conditions permitting, the second priority is to locate the missing person. The third priority is to rescue the missing person or recover the body. The more quickly responders are able to find the missing person, the more likely the operation will be a rescue instead of a recovery. This final priority is represented in the overall GDTA goal of rescuing/recovering the missing person.

The purpose of this thesis is to develop UAV technology to support more efficient WiSAR with less risk exposure to the human responders. Therefore, emphasis in the task analysis is placed on the search plan (goal 3.0) and executing the search plan (goal 4.0) goals. For completeness, a brief overview of the other related goals is provided.

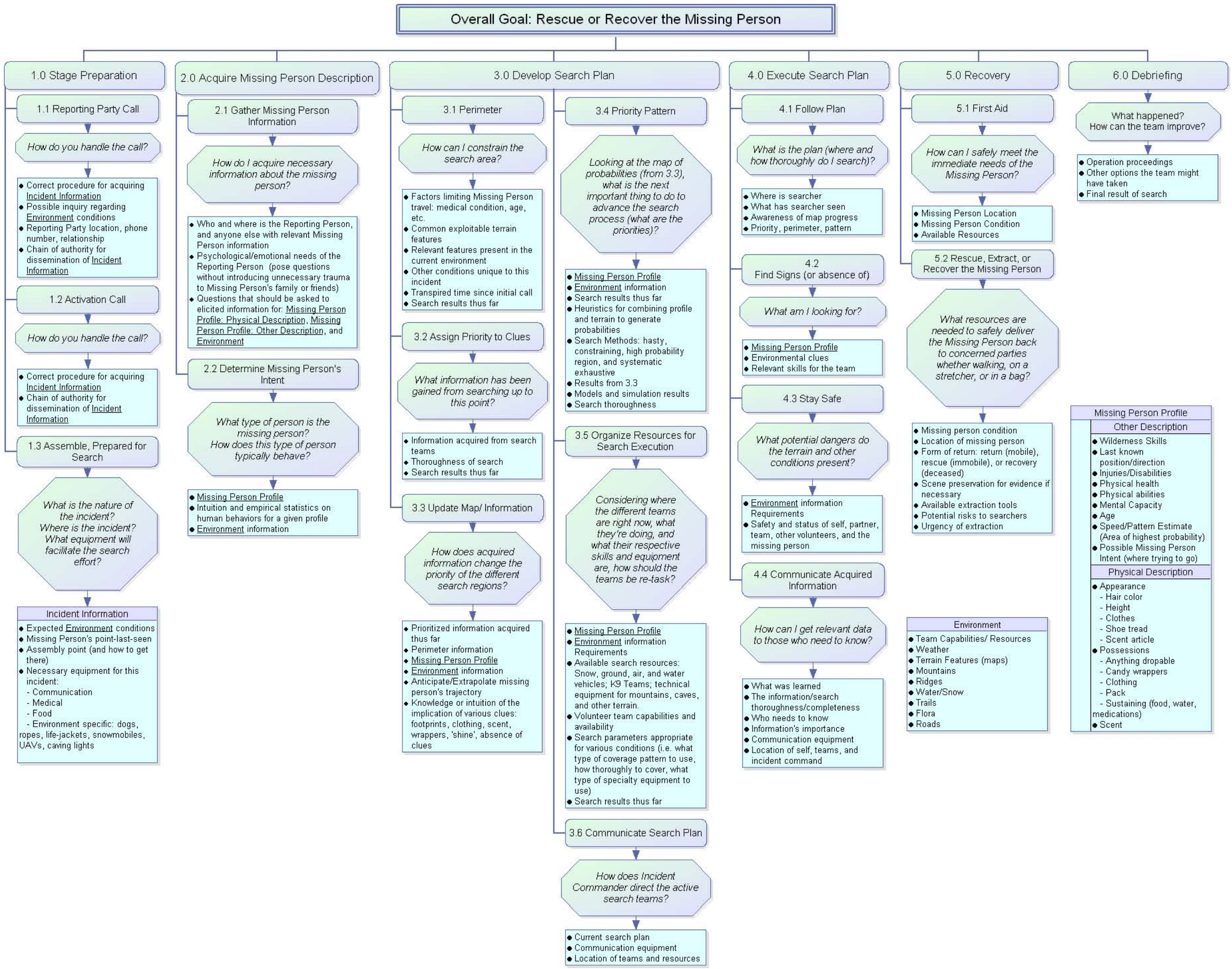


Figure 3.1: The overall WiSAR GDTA results for all high-level goals

3.1.1 Stage Preparation - Goal 1.0

The WiSAR process begins when someone grows concerned over a missing friend or relative. This person, known as the reporting party, contacts the appropriate authorities (such as a 911 call center), as represented by goal 1.0, Stage Preparation, in Figure 3.1. The recipient of the phone call collects the incident information (goal 1.1). The recipient of the phone call attempts to determine from the reporting party where the missing person was last seen, a description of the missing person, and the reporting party's contact information. The call recipient then determines who should be contacted based upon the chain of authority and issues an activation call (goal 1.2).

The WiSAR team, which is primarily composed of volunteers, responds to the call and gathers at a predetermined site and establishes a command center. While first responders assemble, they assess the nature of the incident, where the incident scene is located, potential environmental conditions, and what equipment is required for the response (goal 1.3).

3.1.2 Missing Person Description - Goal 2.0

While the responders are organizing at the assembly point, additional personnel collect the details of the incident (see goal 2.0, *Acquire Missing Person Description*, in Figure 3.1). Authorities contact the reporting party in order to verify the information obtained by the call recipient (goal 2.1). Authorities will also obtain additional information from the reporting party and other relevant individuals (e.g., family and friends) in order to obtain details on the missing person's clothing, appearance, and possessions (goal 2.1) for the missing person profile; see Figure 3.2. Such information is very important in assisting the searchers when analyzing possible sightings and clues. Equally important are the missing person's personality, mental and physical health, intentions, experience with the terrain, last known direction of travel, and any other information that may provide an indication of what the missing person's

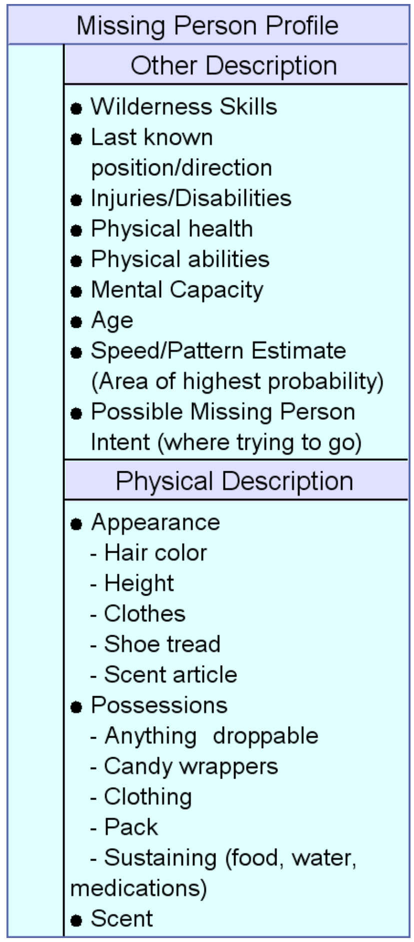


Figure 3.2: The WiSAR GDTA Missing Person Profile information requirements

Figure 3.3: The WiSAR GDTA Environment information requirements

reaction will be in the given situation. This information is employed to develop a missing person profile that is used by the searchers to determine what to look for and where to look.

The incident commander and responders compile their assumptions regarding the missing person's intent (goal 2.2). These assumptions are formulated based upon the developed missing person profile, the environmental conditions (Figure 3.3), intuition, and statistics regarding human behavior. With these assumptions, the WiSAR team begins modeling where to find additional information and planning how to obtain it [49].

3.1.3 Search Plan - Goal 3.0

The third goal for the WiSAR response requires the WiSAR team to develop a prioritized search plan; see goal 3.0, *Develop Search Plan*, in Figure 3.4. The development of the overall search plan incorporates the six subgoals shown in Figure 3.4. The incident commander employs the search plan when determining how to deploy the available resources to perform the actual search.

Establish perimeter - Goal 3.1

The WiSAR team's first objective is to determine, along with the incident commander, the search area perimeter. The intent is to constrain the search area based upon the missing person's profile regarding physical health and limitations, wilderness skills, last known position and direction, and possessions as illustrated in Figure 3.2. Environmental factors (Figure 3.3) such as terrain, weather, etc. will directly feed into the determination of the perimeter. The perimeter decision is also influenced by the time that has transpired since the initial phone call and the search results obtain thus far by family or other concerned parties. The perimeter defines the physical area to be searched.

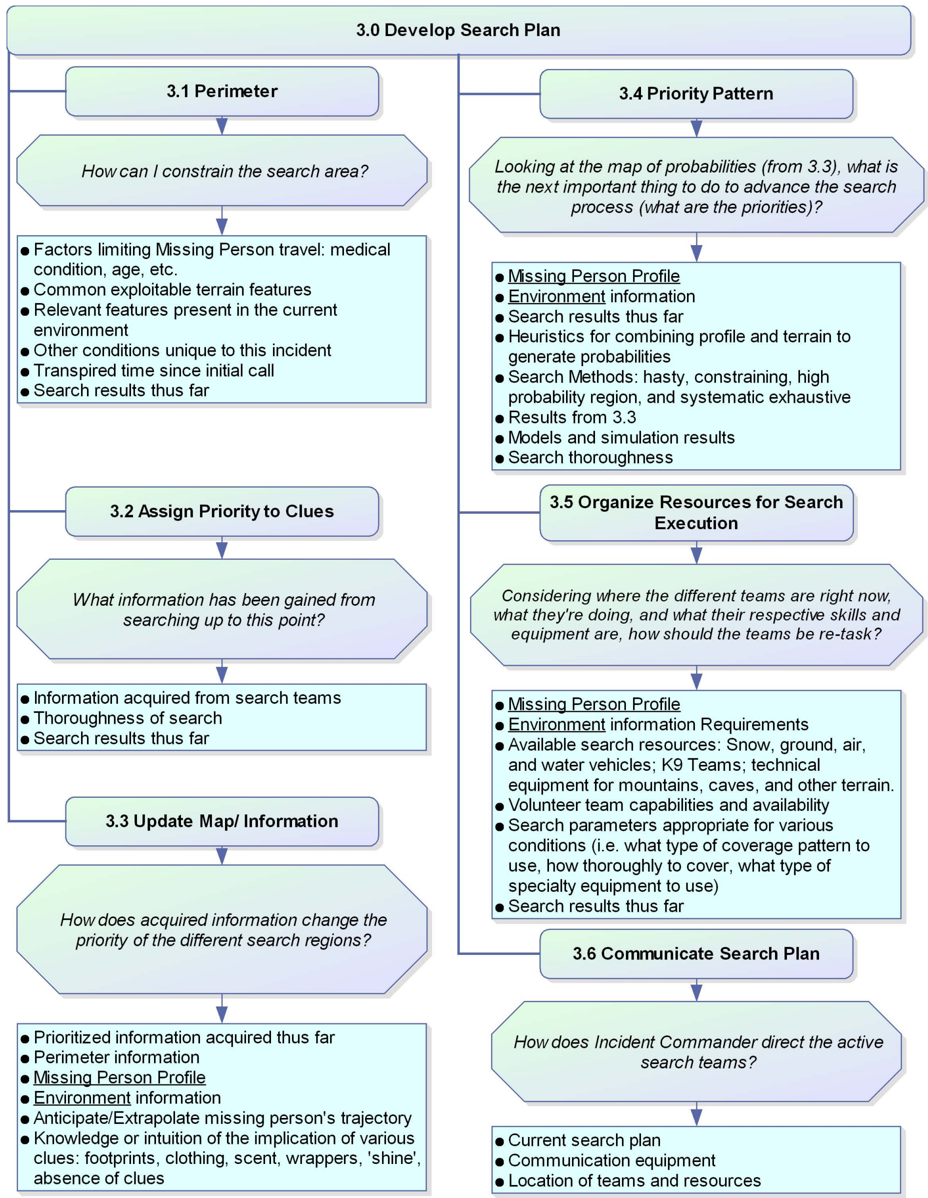


Figure 3.4: The detailed WiSAR GDTA 3.0 goal - Develop Search Plan

Assign priority to clues - Goal 3.2

As information is gathered and the search progresses, priority is assigned to the accumulated information according to its relevance and significance. Since this search is an on-going activity, the assignment of priority to the gathered information assists in determining how the search proceeds.

Update map/information - Goal 3.3

A search map is maintained throughout the process. This map is updated as information is received and evaluated. As search teams cover their assigned areas and report their findings, incident command records the progress of the search. The search map tracks the information accumulated about probable missing person locations. This updated map and information are used to determine the search priority pattern.

Priority pattern - Goal 3.4

The objective of establishing the search priority pattern is to identify the expected value of searching areas within the incident perimeter. The incident commander factors the missing person profile and environmental conditions into a set of heuristics in order to determine probabilities associated with the areas within the search perimeter. An example of such a heuristic is the observation that despondent people tend to seek out high places with a good view of civilization. Probabilities are distributed across the search area in order to guide search plan development. The priority pattern requires consideration of the search thoroughness and results from models and simulations.

Search thoroughness may be represented as the probability of detecting the missing person or an indication of the person's location if such were present in the area. It is necessary to specify the level of thoroughness since dedicating too much time and effort to one area reduces time spent in other areas. A coarse search technique may

be possible if the missing person can hear, see, or call out to searchers (a constraint that is not always satisfied with very old, very young, disabled, or injured missing persons [49]). Similarly, a coarse search may be possible if expected cues are easy to detect, such as bright, discarded clothing.

The priority pattern also establishes what resources should be used and which of several search methods will be employed. Four qualitatively different types of search are used in WiSAR:

- Hasty/heuristic
- Confining
- High probability region
- Exhaustive

Hasty Search. In many cases, the initial model of likely missing person locations has a few regions of particularly high-probability. WiSAR searches often begin with a *hasty search*, rapidly checking areas and directions that offer the highest probability of providing useful information about the missing person. For example, trails, tents, and areas immediately surrounding the missing person's last known location and destination merit hasty inspection. This search is considered "hasty" because the longer the searchers wait, the lower the probability that this type of search strategy will yield useful information. The probability distribution flattens out as time passes and signs fade. The incident commander will often initially employ canine and "man-tracking" teams to follow the missing person's trail. This can be considered part of the hasty search. Additionally, a hasty search can facilitate the execution of subsequent search phases by providing information regarding the missing person's possible location.

Constraining Search. The initial search efforts often include a *constraining search* in addition to the hasty search. The purpose of the constraining search is to find clues

that limit the search area and establish a perimeter for the search. As an example of the constraining search strategy, if there is a natural ridge with only a few passages, searchers will inspect the passages for signs of the missing person in order to restrict their efforts to one side of the ridge or the other. It is important to note that every search that does not provide evidence of the missing person's presence serves to constrain the search by providing evidence of the missing person's absence.

High Probability Region Search. Results from hasty and constraining searches are often used to inform search in *high-probability regions*. As information from these searches and the likely behavior of the missing person become available, the command center divides the search area into sections. These sections are drawn onto maps that are distributed to the searchers as they arrive in order to provide a common language and frame of reference with which to chart the search progress. The incident commander can estimate the probability of finding the missing person in the various sections of the map based upon a combination of experience, intuition, empirical statistics, consensus, and natural barriers [49]. The incident commander then deploys the search teams with the appropriate skills to examine the areas of highest probability. The search teams report their findings as well as an assessment of the thoroughness of coverage as they search an area. The reports allow the incident commander to revise priorities and reassign resources to different areas.

Exhaustive Search. As the high-probability locations are covered, the probability distribution either begins to concentrate on a single region as positive evidence is accumulated, or it spreads out to represent a uniform distribution as negative evidence accumulates for the regions that were initially probable. Eventually, the priority search turns into an *exhaustive search* with the incident commander directing the systematic coverage of a large region using appropriate search patterns. An exhaustive search is typified by "combing" an area wherein searchers form a line and

systematically walk through an area. Exhaustive approaches may produce clues (such as discarded food wrappers or clothing) that indicate the presence of the missing person at some point. If the exhaustive search produces new information, the incident commander may choose to refocus efforts to a form of prioritized search.

Organize resources for search execution - Goal 3.5

For all phases of search, the incident commander and search teams must organize and select the appropriate resources for the present task at hand. The search changes over time based upon search techniques and the information obtained via the search. Using knowledge of the situation, the incident commander selects team members with appropriate skills for a specific step in the search. Likewise, search teams select appropriate skills and equipment to accomplish their portion of the WiSAR goals.

Communicate search plan - Goal 3.6

Once the incident commander determines how to use available resources, the search plan must be communicated to the relevant individuals. The searchers (who may already be actively fulfilling previous instructions) need to know where and how the incident commander wants them to proceed.

3.1.4 Execution of Search Plan - Goal 4.0

The incident commander assigns teams to a particular search technique and search area. The search teams are responsible for executing the search and they have four primary sub-goals (Figure 3.5). The search team is expected to execute the search plan (goal 4.1) while searching for evidence (goal 4.2), ensuring their personal safety (goal 4.3), and communicating their findings (goal 4.4).

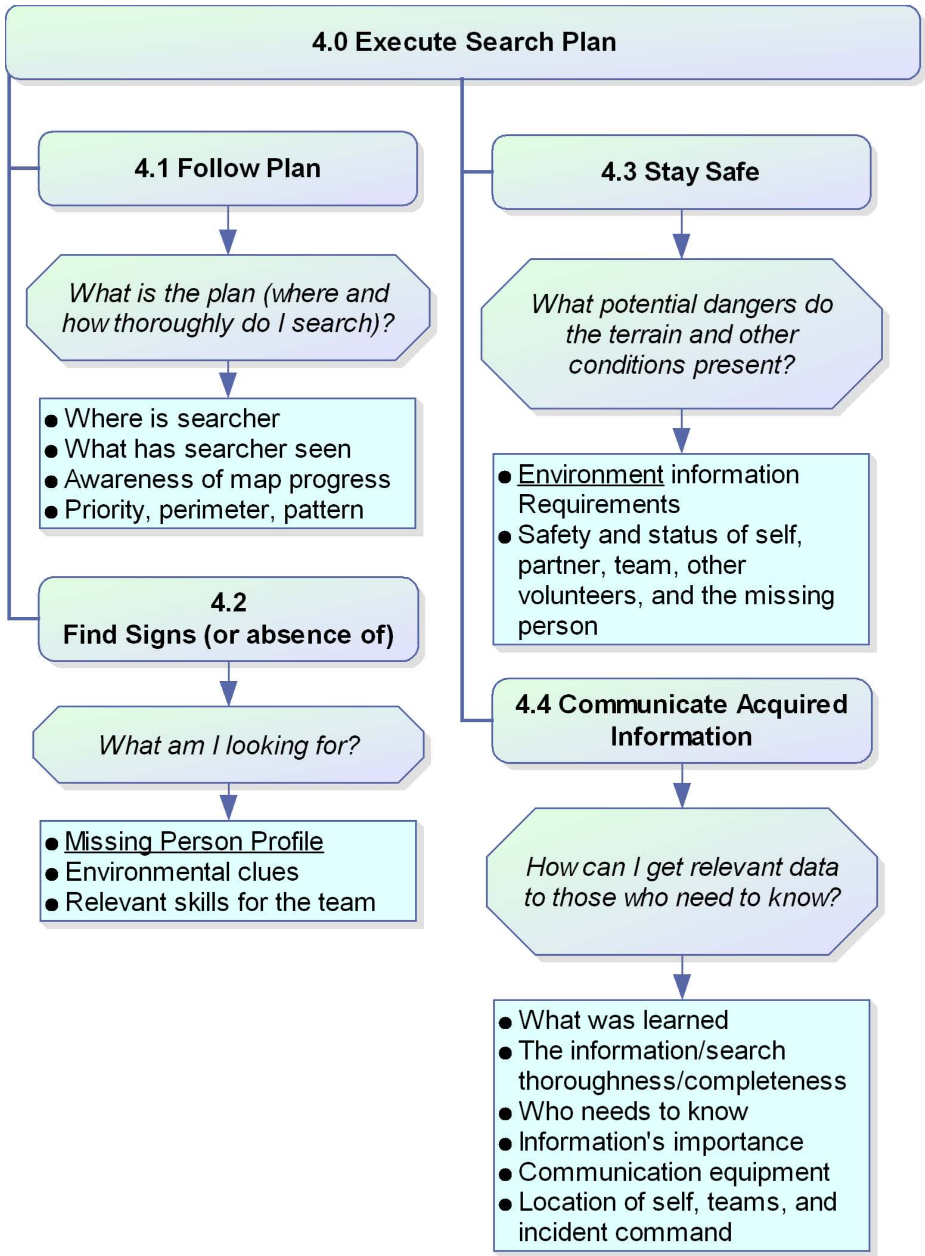


Figure 3.5: The detailed WiSAR GDTA 4.0 goal - Develop Search Plan

Follow Plan - Goal 4.1

Search teams do their best to obtain the information requested by the incident commander. It may be difficult for the search teams to completely satisfy the incident commander's instructions. Environmental elements such as water, weather, vegetation, and rugged terrain may force the searchers to deviate from the planned search.

Find signs - Goal 4.2

Throughout the search process the team looks for evidence (or a lack of evidence), of the missing person's recent presence in the area. The team looks for items the missing person had in his or her possession, footprints, natural or intentional disruption to the environment caused by human passage, etc.

Stay safe - Goal 4.3

Continuously throughout the search process, the search team members' first priority is their own safety. There are a large number of potential hazards to the search team members that they must monitor based upon the environmental conditions and other conditions present in the area.

Communicate acquired information - Goal 4.4

After completing their assigned portion of the plan, each team reports its results to the incident commander. They may also report mid-search as part of a routine update or if some detail warrants immediate attention. The team describes its findings along with its assessment of their significance. When a team finishes searching an area, they will also give an estimate of their thoroughness so that the incident commander knows how likely it is that they missed something.

3.1.5 Recovery - Goal 5.0 and Debriefing - Goal 6.0

The overall GDTA shown Figure 3.1 includes two additional goals representing the recovery of the missing person and a team debriefing. The recovery (goal 5.0) includes administering first aid to the missing person if necessary, followed by the rescue, extraction, or recovery of the missing person. Extraction typically involves technical skill with ropes to remove a person from a precarious location. The rescue involves transporting the missing person to safety. The term "recovery" typically suggests retrieval and transportation of a body. When the search and rescue operations are completed or incident command decides to scale back operations, the team holds a debriefing. The team reviews the incident, the search process, and possible process improvements.

3.2 Information Flow

The GDTA focuses on goals and subgoals in a task together with the information necessary to meet them. However, it does not have a mechanism to communicate the temporal nature of the goals or the flow from one activity to the next. In WiSAR, many of the the tasks are performed simultaneously and information flows rapidly from one task to another. We have extracted the information flow from the GDTA (Figure 3.6) to illustrate how evidence affects the development of the search plan which then influences subsequent efforts to gather evidence.

The search task involves gathering evidence and then utilizing that information to direct further efforts at gathering information. Although it can be argued that concerned parties are already accumulating evidence prior to calling first responders, for the WiSAR team, the information flow begins with the initial details given by the reporting party. Responders immediately consider the urgency of the call based on the potential danger to the missing person and other factors. Combining prior knowledge

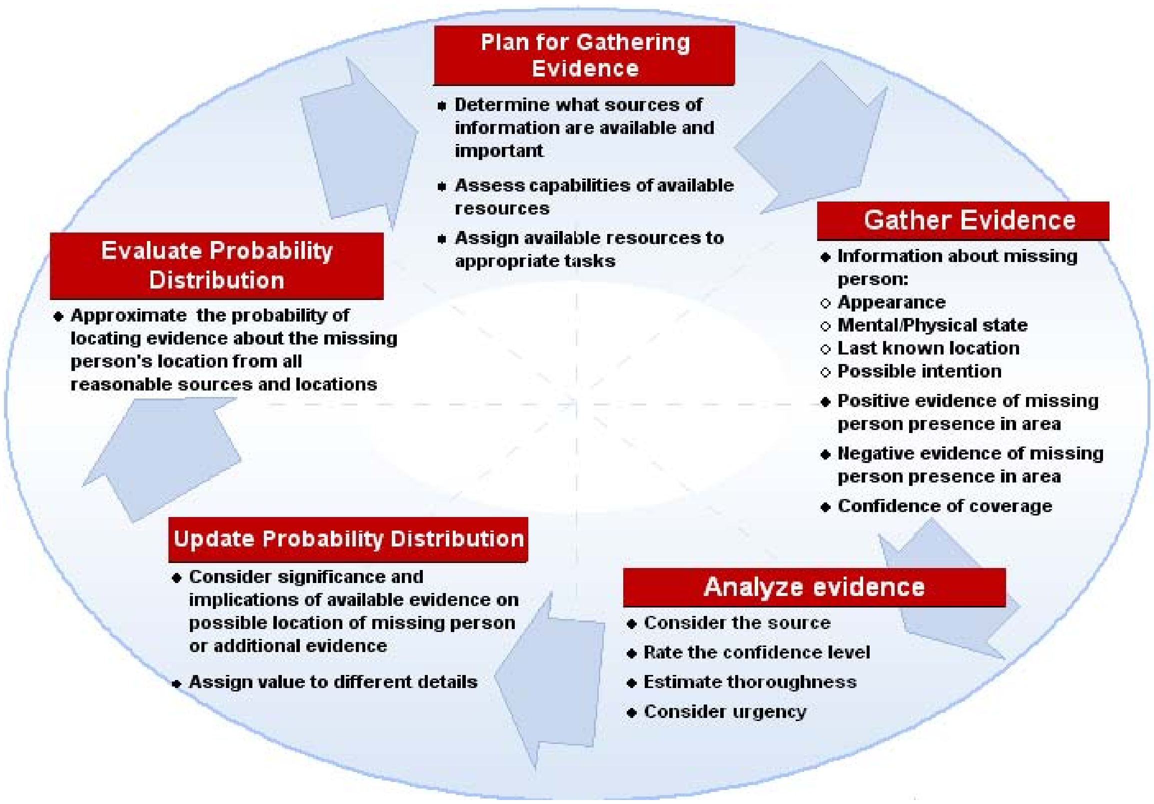


Figure 3.6: Information Flow during Search

and experience with information provided by the reporting party, responders develop an initial model of high-probability sources of additional evidence.

Potential sources of evidence naturally encompass geographic locations surrounding the missing person's point last seen but also include people familiar with the missing person and the missing person's bedroom or other property. After evaluating initial sources of evidence, the WiSAR team develops and executes a plan for acquiring additional evidence. In some cases, this plan may be as simple as waiting to see if the missing person finds himself or herself. In the more interesting case, however, the multiple stages of the information flow are simultaneously active. Different resources are dynamically assigned to accumulating evidence from various information sources as dictated by probability acquiring evidence, usefulness of evidence potentially acquired, risks involved in executing the search, and capability for acquiring evidence from a specific source.

The process continues in parallel as time passes. Time and additional evidence result in adjustments to the probability model of possible sources of evidence which, in turn, leads to changes to the search plan. All evidence affects the expected utility of searching in different areas. The incident commander continually evaluates evidence and redirects available resources in order to maximize the value of the search.

The process may terminate for a number of reasons. Ideally, the WiSAR team locates the missing person (probability distribution converges to a single spike). Work then proceeds on to rescue or recovery. However, the process may also end if the search continues long enough that the probability of the missing person actually being within the search area falls below a certain threshold or if dangers or other constraints (e.g., another incident) cause the relative expected value of continuing the search to fall below a threshold.

3.3 Activity Analysis and Task Breakdown

The introduction of UAV technology into the WiSAR domain must support accomplishing a subset of the goals identified in the GDTA. We anticipate that the UAV will serve primarily to gather information necessary for completing the goals. These information requirements must then be translated into design objectives, such as the following:

- UAV autonomy
- ground control station information presentation for the operator
- procedures required to use the technology
- size and makeup of teams

In this thesis, we emphasize the first three objectives. In this section we emphasize UAV autonomy and suggest some possible procedures for using the resulting technology. We discuss the design of operator interfaces in Chapter 4. Significant portions of this section are the work of Morgan Quigley; see [1] and [19].

3.3.1 UAV-Enabled WiSAR: Task Breakdown

We must consider many different consequences when integrating a new technology into the existing WiSAR process. These consequences include new responsibilities imposed on the searchers, shifts in responsibilities for the searchers, modifications of and integration into existing processes, changes in how information flows, and possible side effects of introducing the technology.

UAV-enabled search is a complex activity requiring closely integrated human interaction with both the operator interfaces and onboard autonomy. Figure 3.7 provides a task-breakdown of UAV-enabled WiSAR. This breakdown was obtained by combining results from the GDTA, observations from field tests, and an activity analysis patterned after the frameworks in [36, 51].

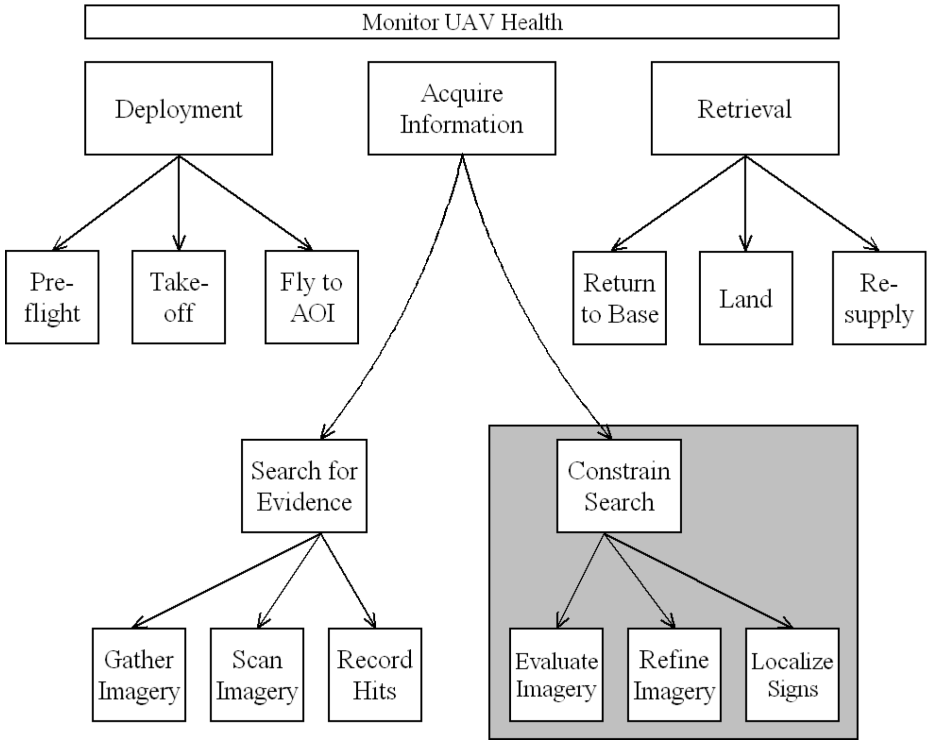


Figure 3.7: Hierarchical task breakdown of UAV-enabled search

This breakdown identifies three new responsibilities for the WiSAR search personnel: monitoring the UAV, deployment of the UAV, and retrieval of the UAV. Maintaining the UAV is a fourth new responsibility, but we omit a discussion of this responsibility in the interest of space.

The task breakdown in Figure 3.7 uses the terms "Search for Evidence" and "Constrain Search" to represent search-related tasks that have been altered by the introduction of UAVs. Sections 3.3.3 and 3.3.4 discuss these two tasks. Prior to doing so, we briefly discuss deployment, retrieval, and monitoring.

3.3.2 Deployment, Monitoring, and Retrieval

When a portion of a task is automated, the responsibility of the human shifts from performing the task to managing the autonomy that performs the task [68]. This shift introduces new responsibilities for the human. The first set of design requirements delineates how these new responsibilities must be performed. These new responsibilities

associated with UAV-enabled search include deploying, retrieving, and monitoring the health of the UAV.

Deployment

The deployment phase commences once the preflight steps are completed. The deployment phase requires the UAV to take-off, climb to cruise altitude, and navigate to the point at which the search is to commence as identified in the GDTA from Section 3.1. For example, the starting point for a hasty search will likely be near the point the missing person was last seen.

Operator Interface. The deployment phase requires that the operator interface support preflight procedures, portray the relationship between the launch point and the search start point, and allow the operator to control travel between the launch and search start point. Preflight steps include checking all sensors and actuators, recording the home base GPS coordinates, and validating the proper setting of control parameters. Finally, the operator selects an initial behavior for the craft.

Autonomy. The initial flight plan typically consists of an autonomous spiral to the selected height above the ground, at which point the UAV enters an autonomous loiter pattern until further instructions are provided [41]. However, the craft could also execute a pre-loaded, fully-scripted flight plan—complete with instructions for obtaining imagery and returning to land at the home point.

Monitoring

Aircraft status anomalies, battery life, and other UAV health information must be efficiently communicated to the UAV operator. Since this information must be monitored throughout all mission phases, Figure 3.7 depicts the monitoring task spanning all other stages.

Operator Interface. The operator interface must allow the operator to confirm nominal behavior and detect anomalies. The operator needs to be able to recognize potential problems on the craft such as electronic malfunction, hardware failure, poor communication signals, and declining battery life. The operator must also be able to track the behavior of the craft to ensure that it is correctly executing instructions and that the correct instructions were issued (because operators are human, we must expect them to make mistakes and the interface should let the operator see what he or she has done and verify that it is what he or she intended). Attention management aides can help operator attend to status information, though this is is a non-trivial problem since warnings and alerts can increase workload and disrupt critical control tasks [4, 48, 63].

Autonomy. The autopilot and ground control station employed in this work include failsafe autonomy modes, which are a form of self monitoring. These are desirable because they can take effect even if communication with the control station are lost or the operator fails to recognize a particular danger. An example of such a failsafe mode occurs when communication with the ground station is lost for an extended time period; under these conditions, the UAV automatically returns to the home base (where communications are likely to be restored or a pilot can assume control via radio control).

Retrieval

Similar to the challenge of deploying the UAV, retrieval is not a trivial task. UAV retrieval requires navigating the UAV to the retrieval point, which may be different from the launch point or home base. The retrieval point during WiSAR may shift locations due to changing weather conditions or discovering a location that better supports communications.

Operator Interface. The key pieces of information required for UAV landing depend on the specific craft structure and capabilities. A craft that requires a runway and careful maneuvering will have different requirements from a craft capable of vertical takeoff and landing. The craft used in this work is sufficiently robust to belly land on the ground without any sort of landing gear. Given the autonomy described in the next paragraph, the operator interface must support the human's ability (a) to identify a landing point and (b) to select an approach vector that is compatible with the terrain and weather conditions. The approach vector is selected such that the approach does not require the UAV to fly through trees or other obstacles. The operator interface should also present the UAV's last known GPS location in case the UAV crashes.

Autonomy. Landing has been addressed in [5, 41]. The UAV automatically flies to a location that is a specified distance from the landing point and then spirals down to a predetermined height above the ground. Upon reaching this height, the UAV breaks out of the spiral and flies the approach vector to the landing point.

3.3.3 Searching for Evidence

The introduction of new technology and the resulting new responsibilities imposed on the operator represent only one consideration. The new technology will also change the nature of how previous responsibilities are performed [67]. Recall that the objective of the search process is to gather evidence regarding where the missing person is or is not located. Without a UAV, this evidence is obtained by ground-based search teams or manned aircraft. With a UAV, locating evidence also occurs through remote video feedback.

The basic steps for a successful UAV-enabled search include (a) aiming the camera to make it likely that visual evidence (either the missing person or some clue about the missing person) appears in the video, and then (b) identifying the evidence's

location in order to guide the rescue team to the missing person. A successful rescue is characterized by rapidly locating a clue toward the missing person's location, since probability of survival drops as time progresses. In the remainder of this section, we use of the generic term "sign" to include any potential clue about the location of the missing person.

Overview

The objective of the searching task during a visual search is to obtain images in which a sign is visible (at least theoretically) by someone viewing the video. This subtask dominates the UAV's flight time and consists of three activities: gathering imagery, scanning imagery, and recording potential signs. The *gather imagery* activity is the fundamental obligation of this subtask and the UAV operator is responsible for directing this subtask. The *record potential signs* activity is necessary to support (a) offline image analysis and (b) localizing the sign for rescue teams. The *scan imagery* activity is not always necessary for completing an exhaustive search, but is necessary if the UAV's trajectory is reactively modified when a potential sign is visible in an image.

Gather Imagery

The gather imagery activity requires the UAV to fly in such a way as to acquire imagery of the search area. Imagery is acquired by planning a path, flying the UAV, and controlling the camera viewpoint to ensure that imagery is obtained of the *complete* search area. The speed and path of the camera's footprint over the ground are the key control variables [32], and the completeness and efficiency of the search are the key performance measures. The path should maximize the probability of locating a sign in the shortest possible time. This task can be simplified by introducing autonomous algorithms that systematically implement the desired search plan.

Scan Imagery

Finding items of interest in the provided imagery is a surprisingly challenging task for an autonomous algorithm. Some search strategies, such as the hasty search strategy, require a human operator to reactively modify the UAVs flight path if a potential sign is found. Such reactive flights require at least a cursory analysis of the imagery so that the operator can view a potential sign, determine the sign's location relative to the UAV, and modify the UAV's path in response. Pixel density, field of view, image stability, and the contrast between sign and background are the key control variables; the key performance variable is the probability of detection given that a sign is in an image.

Record Potential Signs

The UAV operator will make a preliminary classification of the imagery, which will likely include recording potential signs as he or she scans the imagery. This task includes not only saving imagery for a more detailed analysis such as in the localization subtask, but also labeling the imagery with identifying information. This is clearly an action that can be simplified via a well-designed operator interface that allows images and features to be referenced to salient features of the real environment (such as GPS locations or significant landmarks). Potential signs are recorded in world coordinates and can then be employed by ground searchers.

3.3.4 Constrain Search

Constraining the search is an important objective for UAV-enabled search. Finding the missing person effectually constrains the search area to a single point and allows for rescue or recovery, but finding a sign or changing priorities because no evidence is found is also an important constraint. Thus, constraining search includes two basic tasks: localizing a sign, and concluding that there is not sufficient evidence to justify

continued search in a particular area. We will use the generic phrase *locating sign* to indicate both finding a sign as well as concluding that an area does not merit further search. Although automated target recognition technologies exist (see, for example, [43]), we restrict our attention to sign detection performed by the UAV operator.

Overview

Locating a sign with a UAV requires three activities: analyzing imagery, localizing the sign, and refining the imagery, which may require further imagery be acquired. The first two activities are the fundamental obligations of image analysis and the third activity is frequently necessary to validate a clue or localize a sign. Note that the *constrain search* subtask is in a shaded region in the mission hierarchy shown in Figure 3.7. The shading indicates that this task can either be performed simultaneously with sign sensing or be performed at a later time. Note that this task may be performed either by the UAV operator or by a separate "sensor operator" [56].

Analyze Imagery

Imagery can be scanned either in real-time or offline using buffered video. Analyzing imagery with the goal of identifying the missing person's physical location is the primary reason for obtaining imagery; therefore this activity constrains and influences all other activity. The key performance variable for this activity is the probability that a human can detect a sign in an image given a set of image features. This probability is strongly influenced by the way information is obtained and presented. Effective image presentation requires supporting the image analyst's mental reference frames, correlating map and video information sources through techniques such as tethers [37], and employing *a priori* information such as satellite imagery and terrain maps to provide context.

Localize Sign

Once a sign has been identified in an image, it is necessary to estimate the sign's location so that searchers can reach the sign. Estimating the location is often referred to as "geo-referencing" the imagery. If the sign is the missing person, then the searchers must be able to reach the missing person's location in order to complete the rescue. If the sign is a potential clue regarding the missing person's location then searchers may wish to reach the clue in order to determine its relevance and to use it to inform the search process. Much of the sign localization activity can be performed autonomously by employing the UAV's GPS location, the UAV's pose, triangulation, terrain information, and image features [44]. The provided operator interface must permit the operator to identify the sign's features and activate the localization routines. Once a location estimate is obtained, the operator interface must present this information in a coordinate frame that allows searchers to reach the missing person.

Refine Imagery

Image refinement includes techniques that improve the human's capability of identifying the sign, such as stabilizing an image, building a mosaic, orbiting a sign, presenting images in a map context, or obtaining images from different perspectives or at higher resolution [14, 18, 24]. These refinement activities can be classified into two loose categories: enhancing obtained imagery and acquiring additional imagery. Such refinement can be employed (a) to improve the probability that an operator will see the sign, (b) to categorize, prioritize, or discard a sign once a potential sign has been detected, and (c) to improve the estimate of the sign's location. The operator interface capabilities required for this task should allow the operator to request a particular refinement process, such as executing a tracking routine. A reactive flight may require the UAV to fly multiple passes over a sign in order to obtain more images.

The associated operator interface should present information that helps the operator to fly paths that support the image refinement.

3.3.5 Integration into Existing Process

The purpose of introducing a new technology is to simplify the mission, improve mission safety, decrease cost, or speed-up the completion of the mission objective. The mission objective includes many different tasks that often follow a predetermined process. Therefore, it is necessary to identify the existing processes employed during mission execution while specifying how the *new technology integrates into these existing processes*.

The existing WiSAR processes include the procedures used by a search team to locate a missing person. Searches are directed by an incident commander who coordinates the activities of various search teams. Some of these search teams have technical search specialties including medical training, climbing/rapelling, caving, etc. It is likely that UAV-enabled search will require the creation of a new technical search team: the UAV team. How the UAV team interacts with the incident commander and ground searchers is the key question for integrating UAVs into the existing process.

At least three paradigms have emerged in our field tests with members of Utah County Search and Rescue. We will refer to these paradigms as follows: information-only, UAV-led, and ground-led. We now discuss each paradigm. Before doing so, note that UAVs could also be used to provide logistical support in the rescue and recover phase by, for example, scouting paths and entry points through and into rugged areas [14].

Information Only

In the information-only paradigm, the UAV does not directly support a particular ground search team. Rather, the UAV team is assigned an area by the incident com-

mander and then gathers information in this region using, for example, an exhaustive or a priority search plan. The team "covers" the assigned ground, gathers extra information on possible signs, evaluates these signs, and then reports to the incident commander. The incident commander can then dispatch a ground crew to the area if the quality of the information merits.

UAV-Led

In the UAV-led paradigm, the UAV is directly supported by a ground search team. Since the type and quality of information gathered from the air differs from information on the ground, it may be useful to have a ground team available to evaluate a possible sign. In this paradigm, a path is selected for the UAV to travel by, for example, specifying a series of waypoints. The UAV then travels to these waypoints and the ground team also travels to these waypoints. The pace of the UAV search must approximately match the ground crew, which is achievable by having the UAV perform spirals or sweeps around the path. When a potential sign is detected in the video, an approximate GPS location and a description of the sign (either verbal or possibly in the form of an aerial snapshot) is given to the ground crew. The ground crew then finds the location, perhaps with tactical support from the UAV, and evaluates the sign. The information is then either given to the incident commander, or used to refine the path of the UAV.

Ground-Led

By contrast to the UAV-led paradigm in which the UAV occasionally requests information from the ground crew, the roles are reversed in the ground-led paradigm. In this latter paradigm, a hasty search team tries to follow either a scent trail (with dogs) or tracks (with man-tracker specialists). The UAV follows the progress of this hasty search team by flying spirals over them. If the track is lost, the hasty team

can request visual information from ahead, to the side, and from behind the current location of the team. While the ground team is searching, the UAV increases the effectual field of view of the ground team. In this way, the UAV increases the amount of information the ground team can use without corrupting the trail. Importantly, the UAV should probably be flown at an altitude where its sound does not interfere with the ground team's ability to call out and listen for a response from the missing person.

Chapter 4

System Design

The previous chapter contained an analysis of the WiSAR task and a discussion of desirable or necessary features a UAV system should have in order to support the task. This chapter addresses the design and implementation of some of these features in an actual system.

4.1 Platform

The first thing to consider is the physical factor of the system, both craft and control station. WiSAR requires a system that is robust and portable without prohibitive monetary or manpower expense. Furthermore, the time sensitive nature of many searches dictates a system that can be rapidly deployed in wilderness terrain.

4.1.1 Airframe

This research has used a flying-wing type aircraft designed primarily by Nathan Knoebel. The craft (Figure 4.1) has a five foot wingspan and weighs about four pounds. A significant portion of the weight is battery so that the craft has sufficient

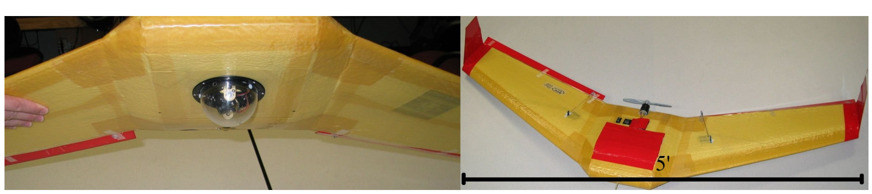


Figure 4.1: Experimental platform

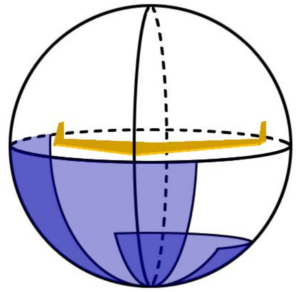


Figure 4.2: Camera gimbal limits

airtime. Another major part of the weight is a Kevlar finish that protects the airframe when landing. A belly-mounted camera is affixed to a gimbal that can point the camera with 135 degrees in the azimuth plane and 115 degrees in the elevation plane. Figure 4.2 illustrates where the camera can point with respect to the craft. The viewable range is biased to the right instead of centered so that it can aim directly out the right wing. Without loss of generality, paths can be planned such that the craft typically turns to the right so that when the craft circles a GPS location, the camera can look out the right wing and remain focused on that point.

The craft design is light enough that an individual can carry it and deploy it by hand. The craft only requires a small clearing to launch or belly land. This makes it possible to rapidly deploy or retrieve the craft even in rough terrain. The craft is controlled by the onboard autopilot discussed in Section 4.2. The autopilot connects to actuators, sensors, and antennae mounted on the craft. Aside from the camera-gimbal, the craft actuators consist of an elevon flap on each wing and an electric push-propeller in the center. The sensors aboard the UAV vary. The craft may carry an infrared camera or optic flow sensors, but the sensor suite always includes inertial measurement sensors to track and control the craft's motion and some way to obtain imagery. A GPS device mounted on one wing supports the control and tracking functionality. The craft sends and receives flight information over a 900 MHz radio connection and transmits video to the ground over a 2.4 GHz link.

Figure 4.3: Radio telemetry link

Figure 4.4: Video antenna

4.1.2 Ground system hardware

The ground control station is somewhat independent from the craft. Hardware on the ground must support the software used to control the craft as well as any necessary physical devices required for communications. The physical ground system must also be capable of supporting any special requirements for deployment and retrieval (e.g., a launch-rail or landing-pad). Beyond these requirements, however, many different hardware setups on the ground could support a given UAV and a particular ground station could control one of any number of different UAVs. The ground station includes a 900 MHz radio modem (Figure 4.3) and an analog video antenna (Figure 4.4). Analog video is digitized by any of a number of commercially available video frame grabbers. The communications antennae can be fitted into a backpack system for portability. The rest of the system is also designed for portability.

Many ground control stations incorporate multiple peripheral display monitors dedicated to different functions (e.g., Figure 2.6). Using multiple display devices increases the amount of available area for visually communicating information and provides intuitive separation for displaying separate chunks of information. However, more pieces typically come with increased expense and can make a system bulky and difficult to transport (particularly on foot). Furthermore, with increased visual area comes increased distance between information elements and increased attention switching costs [62]. This can make it harder to extract information from the system.

In contrast to multi-display systems, we have chosen to focus on a single display system, with preference to a ground system that remains portable even during active use. Requiring a system to be usable while a searcher is walking restricts keyboard and mouse use. A touch screen or other handheld control method may be preferable. Alternatively, the system could be designed to have limited control capability during transportation and then have additional control options if the operator sits down at a temporary base such as a portable table. Although most of our field trials have been on a laptop computer, we have designed software to run on a handheld, touch-based system (see Figures 4.19–4.21). Our intent is to keep the form factor as small and portable as possible without loss of usability in order to meet WiSAR mobility requirements.

4.2 Automation and Abstraction

With an airframe and ground system hardware capable of meeting WiSAR constraints, the next task is to develop the logic and presentation that make the system function and provide the information the operator requires. For many common search operations, the operator should not have to worry about the fact that the video to be searched is provided by a UAV. Ideally, the autonomy and interface will abstract that away, allowing the operator to focus on the tasks of deciding what areas to cover,

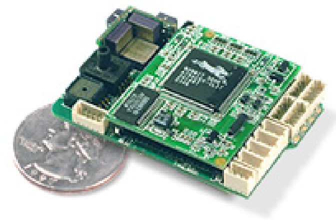


Figure 4.5: The Procerus Technologies Kestrel Autopilot

from what angle, and at what resolution. Once those instructions are provided, the operator can focus more intently on the tasks of interpreting imagery and deciding reactively where to look next—tasks that computers are, as yet, poorly equipped to handle. Although, we cannot yet achieve this ideal on the ground system and must provide some direct control of the craft, we can use automated routines to simplify many tasks and reduce cognitive load on the operator.

4.2.1 Autopilot

The UAV used in this research is controlled by the Kestrel Autopilot (Figure 4.5) originally developed by the MAGICC lab at BYU [11] and marketed by Procerus Technologies. The autopilot is equipped with sensors for measuring altitude and airspeed as well as roll, pitch, and yaw. It also connects to a GPS antenna to determine the craft location. The autopilot transmits this telemetry information to the ground station and also uses it for higher level control of the craft. The autopilot manipulates the different craft actuators to execute commands received from the ground station.

The set of commands provided by the autopilot is relatively simple, but very convenient when compared with direct manipulation of control surfaces. The autopilot can control the camera gimbal to set specific camera angles or point the camera directly at a point in space (within gimbal limitations). By controlling the elevons and propeller, the autopilot manipulates pitch, roll, and airspeed. The autopilot automation builds on these to control heading and altitude. The autopilot can also use these abilities and GPS data to fly to a specific GPS coordinate (waypoint) or circle

a coordinate at a specific altitude and radius (loiter). The autopilot will follow a sequence of waypoints from the ground station, allowing the construction of pre-built search patterns. The autopilot uses a mode system for controlling which of a set of exclusive behaviors to pursue. For example, the craft cannot simultaneously maintain a specific roll angle and a specific heading because one affects the other. Finally, the autopilot provides automated launch and land routines that support the deployment and retrieval steps of UAV-supported search. The ability to stay airborne and follow waypoints partially supports the requirements for getting the craft where it needs to be. The ground station interface is responsible for allowing the operator to specify the necessary waypoint patterns.

4.2.2 Ground station automation

Because the autopilot is developed by another group, we have not had the option to insert WiSAR specific controls. However, additional automation on the ground station can increase system neglect tolerance and provide useful commands needed for WiSAR specific problems. The ground station builds on the command set provided by the autopilot to provide additional commands for the operator.

From the standpoint of the operator, whether the automation logic is on the autopilot or on the ground system makes little difference as long as communications are stable and command execution is tolerant to lag introduced over the communications link. Only some failsafe behaviors that occur when communications decline and time-critical actions (such as holding roll angle) that require very rapid feedback must be calculated onboard the craft. Because the autopilot memory and processing are limited, complex or memory intensive behaviors such as path planning must take place on the ground station. We augment the functionality provided by the autopilot with additional functionality on the ground.

A specific command must be sent to the autopilot to change from one control mode the other; otherwise, the autopilot ignores commands that do not fit the current mode. To simplify control for the operator, our ground station software tracks the different modes and automatically changes the autopilot to the correct mode to execute whatever command the operator attempts to issue. We are currently in the process of adding additional automation and playbook-style behaviors [33] to the ground station software.

As it presently stands, the interface software provides some basic ability to maintain a specific height above ground and automatically fly higher level search patterns. For example, the interface can fly a set of concentric rings (approximating a spiral) for complete coverage of a circular area. The interface also provides a stick-and-carrot control metaphor (Figure 4.6). The "carrot" is an icon that follows the mouse and attracts the UAV. If the UAV reaches the carrot, it first flies over and then circles the point. We implement this control model by sending the UAV to a waypoint where the mouse is pointing. When the craft arrives at that point the interface instructs the UAV to loiter until further notice. When the mouse moves by more than a specified amount, the interface updates the waypoint location and the craft continues to "follow the carrot". A similar model provides control of the camera, allowing the operator to click on a location in the synthetic terrain model and have the video camera point there.

The ground system automation supports WiSAR user interface requirements in several ways. The system design calls for a pre-flight checklist to support deployment. Height above ground functions and ground-based failsafes support the task of monitoring UAV health. The automatic creation of search patterns supports the task of gathering imagery. Other control models assist with the task of refining imagery. Several ground-based functions that process video could be considered part of the system automation because they perform tasks that would otherwise need to be

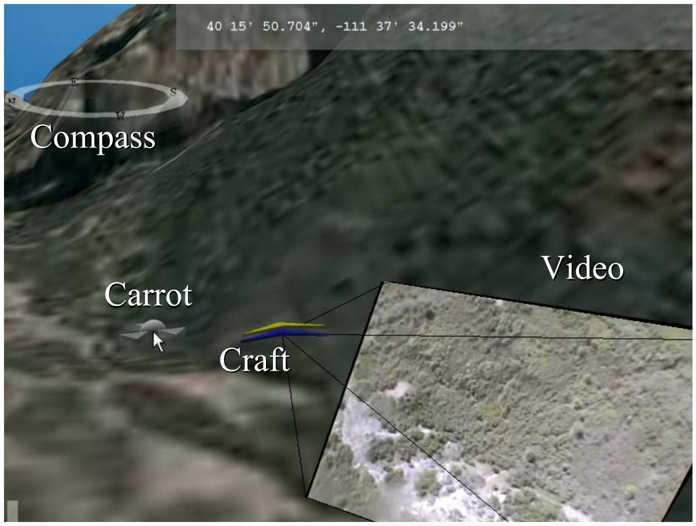


Figure 4.6: Stick-and-carrot control model

done by the operator (e.g., geo-referencing or stabilization), but these are discussed as presentation elements in Section 4.3 because of their visual impact.

4.3 Information presentation

With a collection of available commands provided by system automation, the job of the user interface is to expose those commands to the operator. The interface is also responsible for supporting situation awareness for the operator and presenting information from the UAV sensors to fulfill WiSAR information requirements. Traditional UAV interface presentation methods are not appropriate for WiSAR because they typically require a significant amount of training, may impose a high cognitive load on the operator, and are not designed to support other WiSAR-specific information requirements and constraints. To overcome some of these potential difficulties, we have designed an interface to use an intuitive interaction model and provide the necessary information in an easily understood format.

4.3.1 Intuitive presentation

We have attempted to use an interaction model based on parallel representations [40]. This model uses two similar images to represent a value that the operator can control: a control icon and a feedback icon. The operator issues a command by manipulating the control icon. The parallel image, the feedback icon, shows the current state of the controlled variable. Figure 4.7 shows several common interface elements adapted to use this model. In all but the numeric display, the different modalities are distinguished by color such that the commanded or desired value is represented in yellow and the actual current value is shown in blue.

Our working hypothesis is that this parallel representation supports situation awareness by showing both the commanded and current state in the same frame of reference and by immediately acknowledging the operator's commands. Although we have not formally validated this claim, informal testing suggests that an operator quickly and easily understands the commanded state, the actual state, and the difference between them; visitors seeing the interface for the first time typically need only a moment to understand how to interact with one set of parallel interface elements. Traditional interfaces commonly use one method of input for commands and a completely separate method of output to provide feedback on that command. For example, with a traditional UAV interface the operator may command a change in roll angle by turning a stick or yoke and then rely on an artificial horizon, tilting video, or numeric displays to see the results of the command.

Many common graphical interface elements are represented as a metaphor for something commonly understood so that knowledge transfer from one domain facilitates use of the interface. Many of the common interface elements shown in Figure 4.7 are conceptualized after dials, switches, and gauges commonly encountered as physical control and feedback devices. Metaphors can provide a familiar reference that decreases the need for instructions and simplifies use of the system. Metaphors are

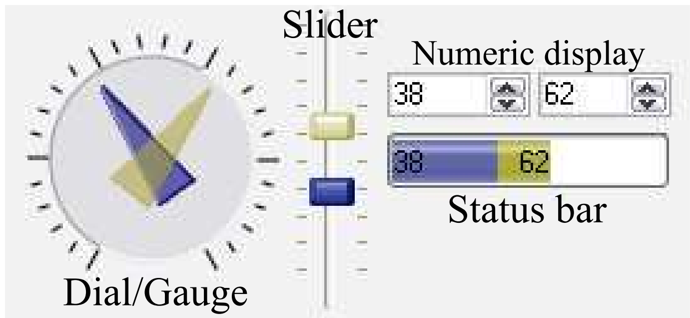


Figure 4.7: Parallel representation with common interface elements

not, however, a panacea to cure all interface difficulties. They must be used carefully and appropriately [6]. In our interface, we have attempted to use iconic metaphors that immediately suggest how to issue specific commands and easily integrate in the operator's mental model.

In traditional flight interfaces, there are typically multiple windows or screen divisions, each dedicated to specific subsystems. These frequently contain numeric displays and analog dials (for example, see Figure 2.7). A numeric input/output (see Figure 4.7) is the most precise method for communicating information, but it may also place the greatest cognitive load on the operator. For example, roll angle can be communicated in terms of exact degrees off of horizontal, but understanding this will require some mental processing to integrate the numeric value into the operator's mental model of what the craft is doing.

As an alternative to numeric displays, an analog dial/gauge representation provides a visible range for comparison rather than numeric values. Thus, analog guages generally come with a slight decrease in precision. However, it is much faster to drag a slider or turn a knob to approximately where it needs to be than it is to type in exactly where it should be.

Analog elements can be combined to provide more sophisticated controls. For example, we combine a slider and a dial into an iconic representation of the craft to communicate and control both altitude and roll. Figure 4.8 highlights this control

icon and the feedback icon. With a touch-screen, commanding a new altitude is as simple as touching the control icon and dragging it higher or lower. Similarly, if the operator wants the craft to turn, he or she simply touches a wing of the control icon and drags it to the desired angle. The feedback icon shows the current altitude and roll of the craft, allowing the operator to track the craft's response to his or her commands.

Figure 4.9 illustrates the speed control and feedback icon. Once again two distinct needles show both the commanded and current values on the same gauge. The operator can use direct manipulation to interact with these different control icons. The operator can then immediately see what he or she has commanded and monitor the progress of the craft as it responds to the command.

Figures 4.8 through 4.10 show the craft and video from a "chase" perspective. The chase perspective is typified by a point of view that follows the craft; see Section 4.3.2. In a chase perspective, direction can be shown using a compass displaced to match the perspective of the terrain (see Figure 4.10). In a chase perspective, the current heading is always forward; so the feedback pointer, which would otherwise be necessary to indicate the current heading, is not shown. The interface only displays the control icon, which can be dragged to a new direction in order to command a new heading.

It is useful to present UAV pose, speed, etc. in a context that supports search. A three-dimensional synthetic environment serves as a suitable metaphor for communicating search-related information. We build a synthetic terrain model using publicly available USGS digital elevation data and satellite imagery or topographic maps. The terrain model is a key interface element. It provides a metaphor for all information dealing with terrain. The operator can annotate areas of the model or plan flight patterns to execute the search. The model can highlight potential dangers presented by the terrain and provide a context for craft and video information.

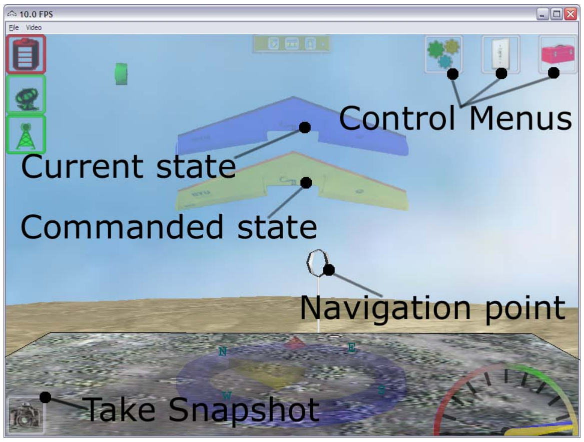


Figure 4.8: Aviation control elements

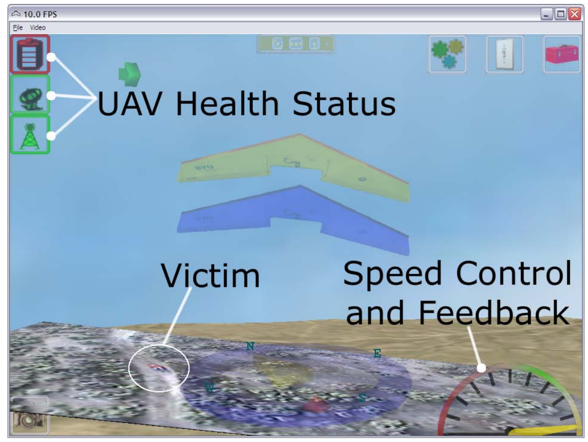


Figure 4.9: More control elements

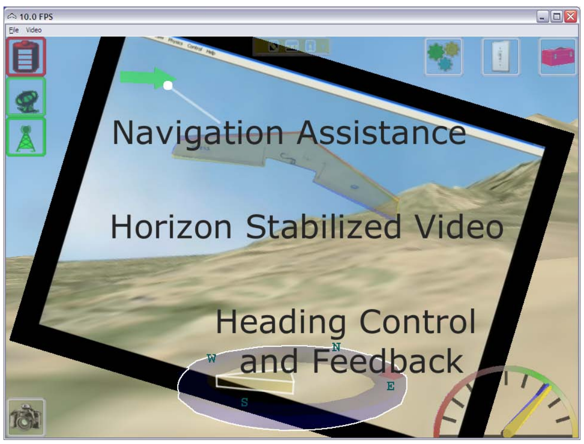


Figure 4.10: Forward facing video

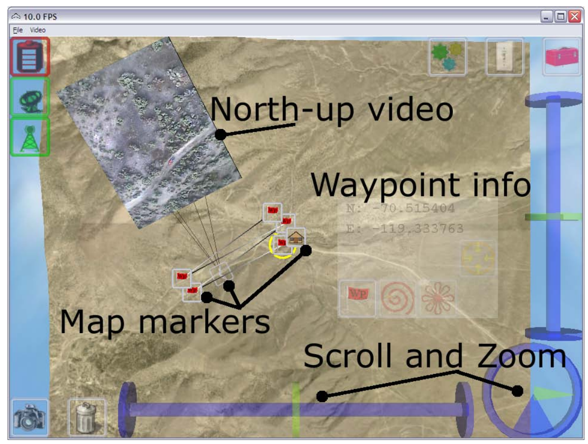


Figure 4.11: Waypoint control

Figures 4.8 through 4.10 show video integrated with the terrain over which the craft is flying. In a chase perspective, the terrain serves to give additional context as discussed in [14]. Using synthetic terrain to provide context, the operator can also point to a location and tell the craft to go there or plot out a complicated flight pattern with a set of waypoints (see Figures 4.11 and 4.22).

In contrast to the chase perspective, Figure 4.11 shows a top-down perspective with the craft flying a set of waypoints and displaying the video off to the side. The video is connected to the craft by tethers. High-detail video is necessary so that the operator can extract information from the imagery. Showing a large area of terrain is desirable because it can provide greater awareness for long term planning. Integrating

the two is important because it helps the operator with the task of *geo-referencing* data extracted from the imagery. Geo-referencing is the process of associating information with physical coordinates (GPS coordinates). Geo-referencing the imagery allows the operator to report information to the incident commander as discussed in Section 3.3.4.

In the case of Figure 4.11, the video is not shown strictly in context as it is in the chase perspective, but is shown at a larger scale and off to the side of the area at which the camera is actually pointing. This is because, at the given scale, the video would be so small as to be unusable. Because it is shown at a larger scale than the terrain, it is difficult to integrate. We draw it off to the side so that the operator can still see and interact with the terrain immediately surrounding the craft. The video rotates appropriately so that north in the video aligns with north in the terrain model. This tether-based solution, while not ideal, may still be helpful. We have explored other approaches [8, 37, 39] but have not yet found a method that satisfactorily communicates high-detail video from a relatively small area of a synthetic environment while simultaneously showing it in context of a large area of the environment.

In a multi-window model with a map in one window and video in a separate window, geo-referencing a feature from the video imagery can be difficult (see Section 5.2) as it may require a complex series of mental transformations to account for craft pose and camera angle. The integrated model simplifies the geo-referencing task by automatically performing these transformations (more accurately than a human operator can), displaying video integrated with the terrain, and providing the coordinates for any location the operator clicks on. This supports the "localize sign" task identified in Section 3.3.4.

In order to take advantage of the automatic geo-referencing, however, the operator must be able to obtain information from the imagery. This requires time

and attention. Bryan Morse and his students are working on integrating the ability to stitch video into a mosaic that the operator can inspect or review whenever time allows. This can greatly improve target detection [18]. At present, in lieu of full geo-referenced mosaicking, we provide a way for the operator to take a video "snapshot" that leaves a geo-referenced copy of the current video frame pasted to the terrain model. This gives additional time for the operator to decide what is in a particular frame of video. This eases the burden on memory and supports the tasks of scanning imagery (Section 3.3.3) and analyzing imagery (Section 3.3.4).

An additional potential benefit of mosaicked video rendered onto the terrain model is that it can be used to communicate what areas have been covered by the video. Such coverage information supports the task of searching for evidence by showing possible holes in the search pattern and allowing the operator to discern the level of detail (and consequently the probability of detection) with which an area has been inspected. In the absence of mosaicked video, the interface provides an estimate of the coverage by drawing a white "smear" from the video footprint. This method approximates the level of search detail by making the coverage smear more transparent when the terrain is farther from the UAV. Figure 4.12 shows the coverage obtained from a spiral search pattern.

4.3.2 Perspective

As discussed in Section 2.4, projecting a three-dimensional synthetic terrain model to a two-dimensional display requires some concept of a *virtual camera*, which defines the frame of reference and perspective from which the model is viewed [7]. The behavior of the virtual camera affects what information is available and how easily it can be understood [64]. For example, if the virtual camera is facing away from a particular part of the model, information from that portion of the model is not available. A top-down virtual camera perspective almost completely obscures the terrain altitude and

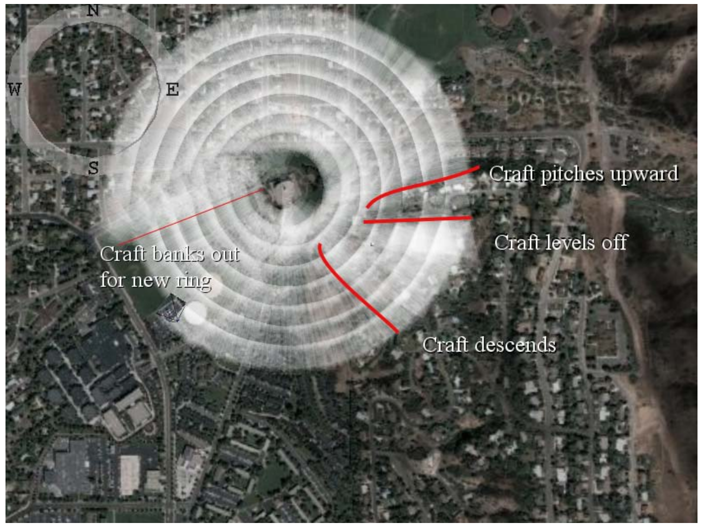


Figure 4.12: Coverage from a spiral pattern

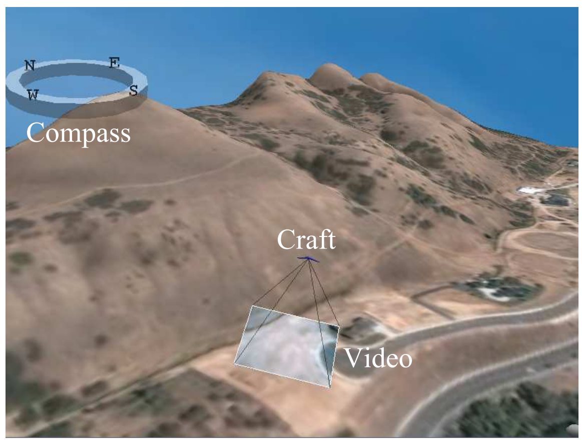


Figure 4.13: Chase perspective

the craft's height above ground, but clearly presents horizontal distances. Different perspectives are desirable for different flying tasks because of the different information they communicate and different cognitive models they support [3, 30]. Some interfaces simultaneously show different perspectives in separate windows [3]. With the single-window ecological model used in our research, we support multiple perspectives by providing a mechanism for changing the perspective when necessary.

One major factor that influences the information available through a given perspective is the *frame of reference* on which the perspective is based. A frame of reference defines the origin and axes of a coordinate system. Perspective is defined

Figure 4.14: North-up perspective

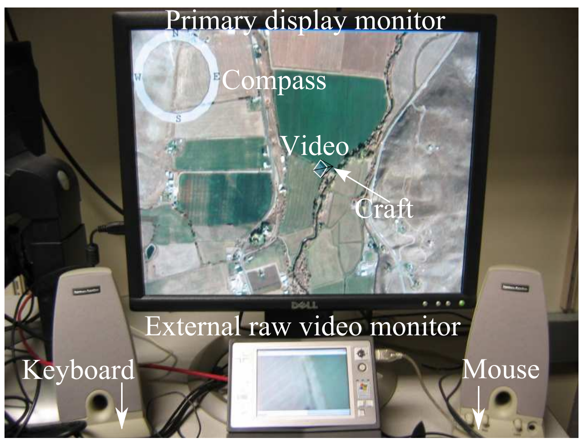


Figure 4.15: Full map perspective

Figure 4.16: Track-up perspective

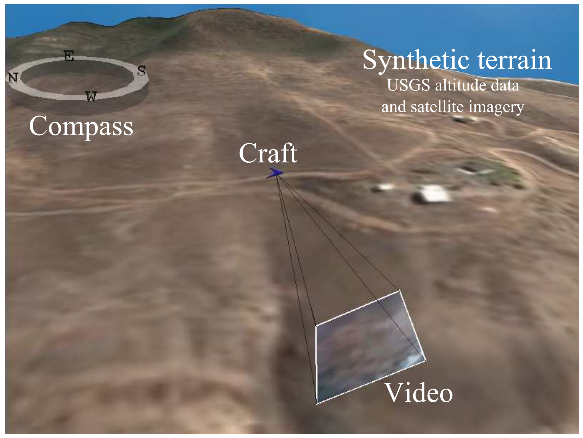


Figure 4.17: Side-on perspective

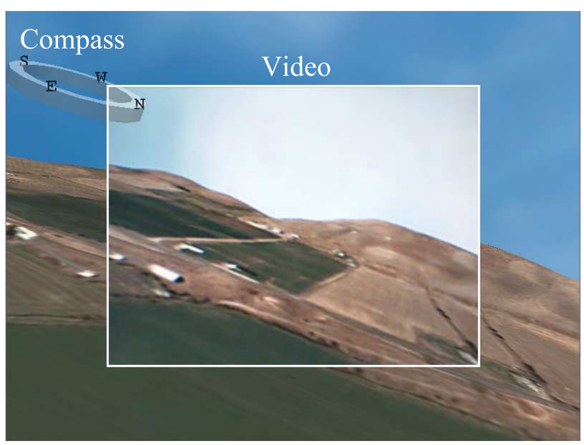


Figure 4.18: Pilot perspective

by an eyepoint and orientation within a given frame of reference. Several different perspectives are common and useful to a UAV control task. *Chase perspective* (Figure 4.13) refers to a frame of reference wherein the origin is defined by the craft, the upward axis is defined by gravity, the forward axis is the defined by the craft's heading, and the third axis is orthogonal to the first two. The eyepoint is behind and perhaps slightly above the craft and is oriented to focus on the craft. A *track-up perspective* (Figure 4.16) uses the same frame of reference, but with the eyepoint looking down from above the craft so that the direction the craft is flying is upward on the display so that when the craft turns, the visual effect is the terrain rotating in the opposite direction. A *north-up perspective* (Figure 4.14) still uses a coordinate system centered on the craft, but all three axes are defined with respect to the terrain: up, north, and east. As with the track-up, the eyepoint is looking down on the craft, but the craft turns within the display and the terrain remains in a constant, north-up orientation. A *full-map perspective* (Figure 4.15) uses the same axes as north-up, but defines the coordinate system with respect to the terrain instead of the craft. The eyepoint is, once again, looking downward, but this time from a sufficiently high vantage point to see all or most of the relevant search area. These perspectives can be contrasted with a *pilot's perspective* (Figure 4.18) that uses a frame of reference built completely around the craft (i.e., one axis aligned with the wing, one axis through the top of the craft, and one axis through the nose). The eyepoint is located in the craft and looks out the nose.

As an aside, the pilot's perspective differs from the others because it does not use a gravitational reference. At first glance, many people interpret Figure 4.18 to be showing video from a craft banking to the right when, in fact, it is banking to the left. Of course, a photograph does not communicate the same optic flow that comes from live video, but the image still serves to illustrate the potential confusion that

faces a ground-based UAV operator without pilot training trying to interpret data through a pilot's perspective.

In our interface, the virtual camera that controls the interface perspective functions by keeping track of two points: the eyepoint and the focus point. These two points can be defined with respect to the terrain, the craft, or the video. For example a point can be defined as being 20 meters behind or to the side of the craft or 20 meters south of the craft regardless of where the craft turns. When the operator wishes to use a specific perspective, he or she may select a given perspective from a pre-configured menu of useful perspectives such as those described above. The operator may also directly manipulate the virtual camera as necessary to obtain a specific vantage point. The ability to change perspectives ensures that the operator can get whatever information is available in the synthetic environment, though not necessarily in a timely manner.

Changing perspectives can potentially confuse the operator. In particular, we have some preliminary anecdotal evidence that a large instantaneous perspective change is disorienting and may temporarily affect situation awareness negatively (see Section 5.1.2). To avoid this we use a quick but smooth transition from one perspective to another. There are many different ways for smoothly transitioning between perspectives. The simplest method is to move the eyepoint and focus point linearly from their current positions to their intended positions. Other, more cinematic methods may look more impressive, but looking better is not necessarily more effective at supporting situation awareness [53].

We believe that a ground-based UAV operator without pilot training may understand rotation in the horizontal (azimuth) plane differently than rotation upward or downward (elevation) such that a perspective rotation through both may be confusing. This hypothesis still requires formal validation. The current virtual camera transition model only rotates in one axis at a time. If the interface were using a chase

perspective with a craft flying southeast and the operator decided to use a north-up, map perspective, the virtual camera would tilt downward while shifting upward. Once the virtual camera was facing downward, it would rotate counter-clockwise (the shortest direction of rotation from southeast) until north was aligned upward on the display.

4.3.3 Attention

In the WiSAR domain, we cannot guarantee that the operator's full attention is centered on the interface. However, when the operator does focus on the UAV control interface, we need to make the interaction efficient and productive. There are a number of information items competing for attention. It is important to control the information presented so that the operator is not distracted by unnecessary elements. Since the operator may nevertheless be distracted by responsibilities outside of the interface, it is desirable to have easily understood information available when the operator does pay attention to the interface. Although the ideal interface presentation will vary based on operator habits and intent, we can use known attention management and information organization techniques to present important information, while simultaneously minimizing clutter, confusion, and distraction.

The first technique is to use transparency to decrease the salience of certain interface elements but keep them usable. Harrison et al. have explored the use of transparency in interfaces and found that there is a trade-off [25]. If an information element is too transparent, it might as well not be there; it is nearly impossible to find, decipher, and use and only serves to obscure whatever it overlaps. If an information element is too opaque, it completely covers what is behind it and negates the benefits of the transparency. However, careful use of transparency can improve use of the interface. We use interface elements with variable transparency.

As mentioned previously, information obtained through video imagery is the primary interest; controlling the craft is auxiliary to that. The small display area and the integrated paradigm force us to frequently overlap interface elements. Consequently, our design makes most interface icons transparent until they are needed. This use of transparency keeps information available but unobtrusive so the operator can focus on the search task. It may take a little longer to find these transparent icons, but the additional time is small and the benefit is added functional area [23]. For example, an icon communicating approximate battery strength with a status bar can sit unobtrusively transparent off to the side, giving information but also showing terrain underneath (Figure 4.9).

A second technique for managing attention is to present extra information when an operator interacts with an icon. For example, touching the battery icon can turn the icon opaque to acknowledge the operator's action and also cause the interface to provide additional battery information such as the exact (numeric) voltage and estimated remaining flight time. Another example is to use menus for rarely used interface elements. Infrequently used interface elements that are not time critical can be completely transparent until the operator clicks a menu icon at which point they become temporarily visible to acknowledge operator action and provide additional functionality or information. Afterward the icons then fade away. This also reduces clutter.

A third technique for managing attention is to change icon salience when a particular information element needs attention. For example, when battery power or the communications signal fall below a certain threshold, the interface may attempt to attract the operator's attention by changing the relevant icon's color, opacity, or size. It may also use an animation such as flashing or swelling to draw attention. Currently, the appropriate icon turns red and becomes more opaque when something happens that requires attention (e.g., low battery or a faulty communication link).

We use this behavior in an attempt to attract attention without being an annoyance (and because it is easy to program).

A final technique for managing attention is to use audio or haptic (e.g., vibrating) signals to communicate certain messages. These non-visual signals can be advantageous in some situations because they can alert the operator even if he or she is looking away and also because information through non-visual media may be easier to handle than additional visual information in a task that is already visually demanding [21]. The interface uses some simple audio acknowledgments but may benefit from presenting additional information through alternative channels. On the other hand, a WiSAR volunteer may already be using his or her audio channel to its limit to communicate with other team members and the incident commander.

We currently use simple versions of these attention management and information organization techniques. Their validation as part of the interface remains for future work. These elements are designed to add increased support for the tasks of monitoring the UAV and gathering imagery without detracting significantly from the tasks of scanning and evaluating imagery.

4.4 Interface evolution

The software we have used for testing and validation has gone through several incarnations and development cycles and will likely need to go through several more before the technology can be used in genuine missions. The initial model (Figure 4.19) was a proof of concept interface developed by Morgan Quigley to show that it is possible to control the UAV with a handheld device [40]. The software ran on a PDA that used a simplified command set to send and receive information through more sophisticated software running on a laptop computer. The interface displayed and allowed the operator to control altitude as well as roll or heading, automatically putting the craft autopilot into the correct mode to execute the given command. An operator using

Figure 4.19: Original handheld interface

this interface could successfully control the craft, but without video or any location information, flight was only feasible if the craft was visible to the operator.

Using Quigley's original concept, we created another system that also ran on a PDA but was independent of any other software and consequently more portable. The system required only a radio modem and video antenna (Figure 4.20). This system incorporated the ability to commanded altitude, roll, and heading by dragging a model of the craft. Dragging on either wing sent a command to roll while dragging the center of the model changed the command altitude of the craft. Two models, one yellow and one blue, served to show both the commanded state and the current actual state of the craft. The controls for this system were displayed as transparent icons overlaid on the video. Video filled the entire (rather small) display. This system also incorporated a geo-referenced map that could be called up to see the location of the craft or plot waypoints. Because of screen size limitations, the video and terrain information could not be displayed simultaneously.

For several reasons, we redesigned the software to run on a more sophisticated device (Figure 4.21). By moving the interface software to a handheld computer with a more powerful processor, we gained valuable display area and also gained the abil-

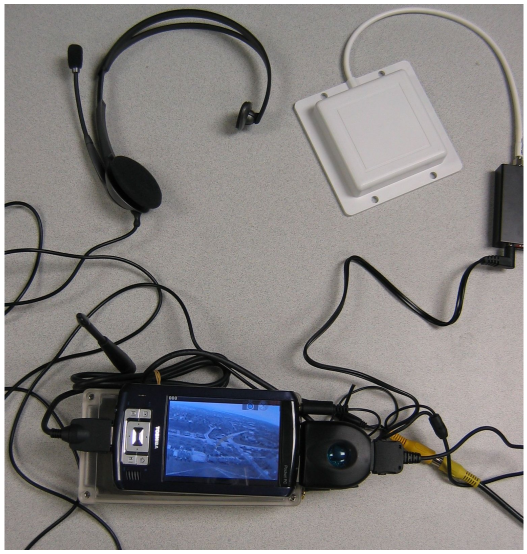


Figure 4.20: Handheld PDA setup with video

Figure 4.21: Vaio handheld interface

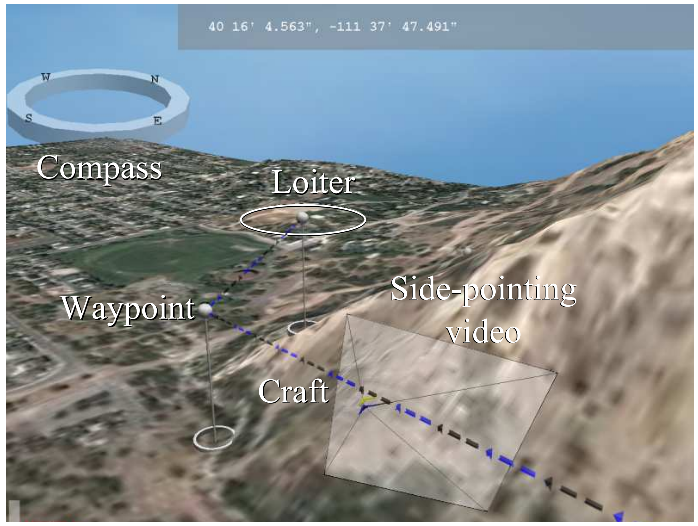


Figure 4.22: Full 3D interface—following waypoints

ity to execute more complicated commands. One important feature that was not available on the PDA was a 3D graphics card. A graphics card allows the interface to display three-dimensional terrain data with integrated video. With this version of the software, we incorporated a drop-down menu system and many different control options and icons. We provided two different perspectives for two different types of control: chase perspective for near-term control (roll or heading) and map perspective for long-term control (waypoints). We also began to use the concept of a movable virtual camera to generate the different perspectives and provide smooth transitions between the two perspectives.

Because of how the software for the handheld was originally constructed, we could not run certain experiments that we wanted to explore. In particular, the design only provided the chase and map perspectives and the operator could only interact with the synthetic terrain (to place waypoints, for example) from the map perspective. We designed and developed another version of the software that uses the architecture described in Section 4.5. The new design (Figure 4.22) allows for arbitrary perspectives and allows interaction with the 3D terrain regardless of perspective. This means that the operator can annotate or place waypoints on any terrain the virtual camera

can see. These features were important for the experiments described in Chapter 5. At the time, this version of the interface is under active development. We are adding new features as more research becomes available and integrating some elements from previous versions of the software. As it currently stands, the software is an effective research tool and we expect that many elements from the design will eventually be incorporated into a production version that will be used by first responders in the field.

4.5 Software architecture

With well-designed software, useful elements and ideas are more easily adapted, updated, or incorporated into other software projects. Good design simplifies the process of developing a final product. We have learned several lessons about applying principles of software engineering and interface design for a UAV interface utilizing a synthetic environment. As always, a modular approach is important for creating flexible and maintainable software. In this case, it is especially important because the software is designed to be used both for laboratory user studies in simulation and field trials. In this section, we present our current software design and the reasoning behind certain design decisions.

The primary requirements that we must satisfy are structuring the software so that (a) information flows where it needs to be, (b) the code is easily maintained, and (c) it can be used for both field and laboratory experiments. Figure 4.23 shows the high-level structure and flow of the code. Inputs are telemetry data and video from the craft as well as actions from the operator (e.g., keyboard, mouse, microphone). Outputs are commands sent to the craft through the radio link and information sent to the operator through the display, audio, haptics, etc.

The interface currently connects to the radio modem link through a serial connection. The simulator can communicate over a serial or TCP/IP link. When

working with the simulator, the interface can open a separate TCP/IP connection and send scripted commands to load targets, change maps, or launch the craft. During flight, the software records all telemetry to a file, which can later be loaded and replayed complete with video. Separate modules represent each of these methods for getting data to the interface. These modules can be easily interchanged or another can be created as necessary without affecting the rest of the software.

Once data is received from the craft or ready to be sent to the craft, it must be converted from or to the format that the autopilot understands. The autopilot used on the craft is still under active development and the API for communicating with it changes occasionally. To simplify communications between the interface and autopilot, a translation layer transforms information flowing between the craft model and the communication link. When telemetry data is received, the translation layer interprets the data and informs the craft model of the latest craft state. When the craft model has a new desired heading or altitude, it sends the necessary command through the translation layer, which formats the request appropriately for the current autopilot configuration. The most frequently changed variables for communicating with the autopilot are loaded from a simple file that enumerates the identification numbers required for the different available commands. More complex autopilot modifications occasionally require changing the translation layer but do not affect the rest of the code. With this design, if we were to use another autopilot with a completely different API but similar abilities, only one small portion of the code would need to be changed and the interface would function the same.

Built on similar principles to the telemetry input and translation layers, using video requires one module that handles acquisition of the imagery and keeps an image buffer filled with the latest frame. Our software uses freely available libraries (DirectShow and OpenCV) to capture live video and to load video from file when replaying a saved flight. When a frame is acquired, a separate module stabilizes

and enhances the image using code written by Damon Gerhardt [18] and Nathan Rasmussen. Each video object is associated with a craft model that understands how the video stream should be displayed given the craft pose and camera angles.

The craft model is central to the control interface. It is a software representation of the current and desired craft states. This software object presents methods for everything the craft is capable of accomplishing. With this design, the interface can have multiple ways for an operator to issue a particular command (mouse, keyboard, audio, etc.) that all access the same method. When the operator issues a command, the automation logic compares the current craft state to the commanded state and issues whatever commands are necessary to execute the command.

The state prediction portion of the craft model is currently only partially implemented, but it is intended to serve many purposes. Communications between the craft and the ground station introduce a certain amount of lag, which can make controlling the craft difficult. State prediction can "quicken" the interface and show the operator a good guess of what the craft is currently doing and thus facilitate control [29]. Prediction can also support certain automatic behaviors such as height-above-ground maintenance. By looking into the future a few seconds, the automation can determine whether or not the current course of action is safe and improve neglect tolerance by taking action if it is not. Finally, accurate prediction can support situation awareness by providing a way to show what the craft will do within a certain time window and how that will be different if the operator issues a certain command (such as the tunnel-in-the sky display [34]).

The terrain model holds information about the area of operation. This object encapsulates terrain height information as well as geo-referenced imagery. Multiple images can be associated with an area: satellite photos, topographic maps, etc. Other information associated with the terrain (such as search patterns, video coverage, and area annotation) is also logically part of this module. Information flows between the

terrain and craft models, allowing the craft to monitor height above ground or follow a set of waypoints and keep track of what areas have already been visited.

These models are unified by the display and event logic modules. The display module handles the logic of communicating the information stored in other modules to the operator and the event module handles information received from the operator. Encoded in the display logic is how to format information for the operator and when to show different information elements (e.g., icons, menus, etc.). An important subcomponent of the display logic is the virtual camera, which determines the perspective and frame of reference used to graphically communicate 3D information such as the terrain and craft state. The event logic provides a pathway for information to flow from the operator to the software system. It handles mouse movement and key presses and examines the interface state to determine what should happen as a result. This object exercises influence over almost all other objects, changing states and issuing commands in response to operator actions. Arrows exiting the event logic module are omitted in Figure 4.23 for simplicity and visibility.

The final high-level component in the system architecture is a script module. When running controlled experiments, we often want scripted events, such as automatic changes in perspective, to take place. The script module can be configured to keep track of an experiment and modify the behavior of the interface according to the independent and dependent variables of interest. Typically, controlled experiments are run in simulation; the script module can also attach to the simulator and launch the craft or load a new terrain model as necessary. With some simple configuring, this module allows us to validate specific interface features.

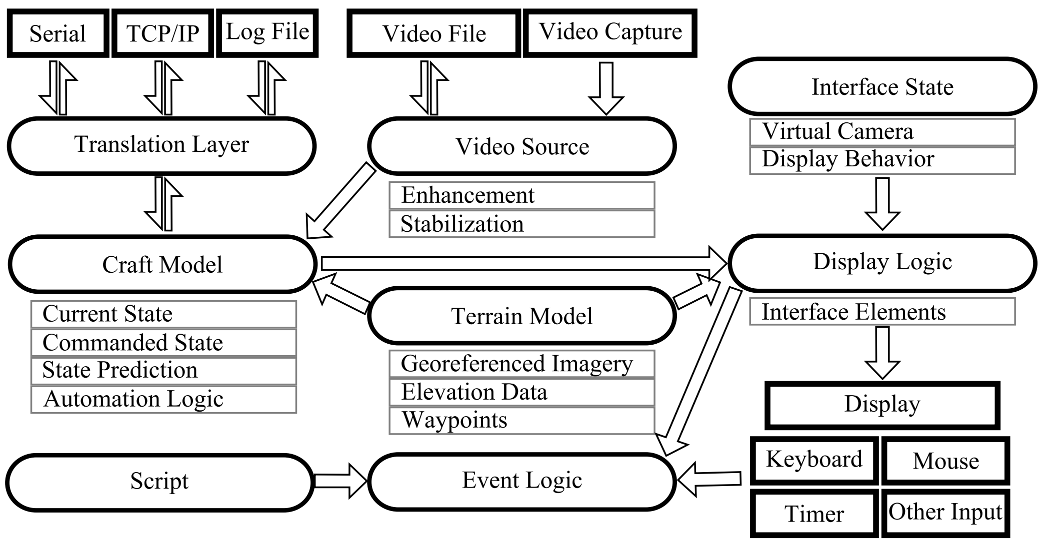


Figure 4.23: Ground station software architecture

Chapter 5

Validation

A large number of design decisions go into making a UAV control interface. These decisions affect the usability of the interface in various ways separately and may also have higher-order effects when combined together. Validation of all features and their combined effects in a full interface system becomes a combinatorial impossibility. In creating the interface described in this thesis, we have made an effort to make design decisions according to general interface principles and related research. We have also tested some features through controlled experiments and partially structured field studies. In this chapter we discuss some of the work we have done to experimentally validate the interface design along with practical justification for other design decisions.

5.1 Small-scale experiments

Prior to running full scale experiments with this interface, we conducted a few small preliminary tests. These studies used only a small number of subjects because the data demonstrated overwhelmingly strong results. Two of these studies are described in this section.

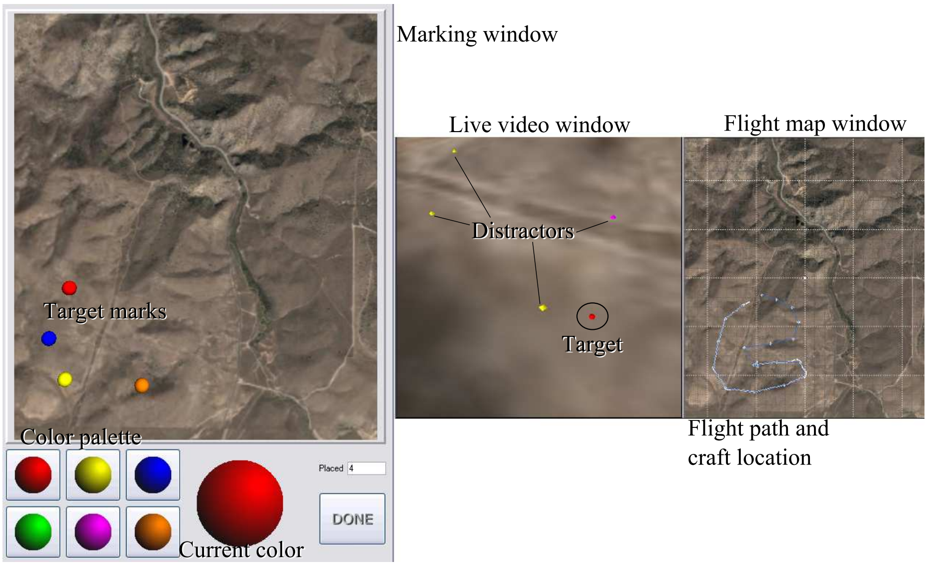


Figure 5.1: Multi-window detection study

5.1.1 Multi-window target detection study

The experimental setup of the first study is shown in Figure 5.1. Five unbiased participants observed four, five-minute flights. The flight map window (on the right) shows a map with the full flight path and current location of the UAV. As the craft flies over terrain, the live video window (center) shows video received by the UAV camera. The marking window (left) provides a map of the same location as the flight map window and allows test participants to mark locations with colored spheres.

This experiment took place on a 19 inch LCD monitor. Participants used a regular optical mouse to complete the task and were paid $10 for their participation. When all four trials were completed, the participants filled out a brief subjective survey on their experience. All participants reported normal or corrected-to-normal vision.

During the four flights, four different video presentations were shown to participants in random, counter-balanced order: downward, downward-stabilized, forward, forward-stabilized. The downward trials simulated a camera pointing directly out of the bottom of the craft. The forward trials actually used a camera at a forty-

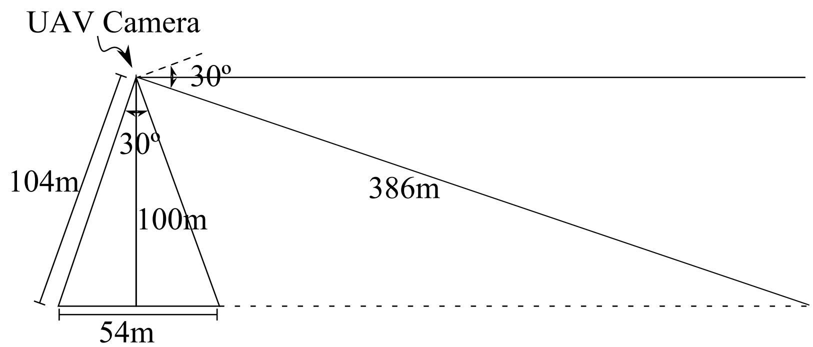


Figure 5.2: Straight forward and straight downward video

five degree angle. We did not use directly forward facing video because, for a given flight path, a straight forward facing camera sees a completely different area than a straight downward facing camera, making it difficult to keep the flight paths and target distributions consistent across the experimental conditions. Furthermore, in video from straight forward facing video over flat terrain, targets appear, at most, about one-fourth as large in the video, making it quite difficult to distinguish targets from distractors (ground targets in video are at least four times farther away; see Figure 5.2). The stabilized trials maintained the camera in a constant angle between the craft and the ground even when the craft was turning. The non-stabilized trials kept the camera fixed with respect to the craft so that when the UAV turned one direction, the video footprint extended in the other direction.

The experimental task required participants to integrate information from all three windows by first recognizing a target (spheres) in the video window while ignoring distractor artifacts (pyramids) and redundant sightings of the targets. After identifying a target, the participant was to look at the map to see where the craft was. From that, the participant could deduce the location of the sphere. After selecting the matching color from the color palette at the bottom of the marking window, the participant marked the location of the target on the marking window map.

We hypothesized that participants would encounter greater difficulty in completing the task with the unstabilized video because the video swings around whenever

the craft turns or changes altitude which may be disorienting. We further hypothesized that a downward facing camera would support greater accuracy in marking target location because targets are directly under the craft, but that the forward facing camera, with its larger video footprint and longer time-in-view would support identifying more targets.

Unfortunately, we found that the ordering and the four different flight paths introduced very strong confounding factors. Because the design lacked an initial practice phase, the first trial always went badly as participants became accustomed to the task. The different flight paths covered approximately the same distance but covered very different terrain and followed very different courses. On some paths the craft looped back on its path while others covered unique areas. Some had sharp turns while others had more gradual turns. These factors had such a confounding effect on the data that we stopped the experiment to begin working on a different design.

In this experiment, we found that the participants marked a high number of redundant targets. There were only 10 targets visible in each flight, but the participants marked, on average, 16.35 targets per flight. Three of the five participants commented on the difficulty in discerning redundant targets. One strong contributing factor to this redundancy was that the participants' attention was stretched across the three different windows. The experiment described in Section 5.2 supports this conclusion.

As a secondary observation, during all five cases we focused a camera on the participant's face to observe where he or she focused. We found that participants split their attention fairly uniformly across the three different windows, spending one to five seconds on each window. They typically followed a consistent pattern of jumping from one window to the next and then occasionally sitting forward and paying more attention to the live video window. From this jumping pattern, we expect a significant

cognitive load on the participants as they attempt to gather, remember, and integrate information from the three windows into their own mental model of the situation.

5.1.2 Path recall study

Another study attempted to track the effect of different perspectives and perspective transitions on target detection and flight path recall. As in the previous study, the craft flew a preprogrammed course and the test participants observed the flight without controlling it. Eight unbiased subjects participated in this study. The task was similar to the previous study, but in this experiment the participant observed a flight from a third person perspective using the synthetic environment interface framework described previously. Throughout each flight, targets (spheres) and distractors (pyramids) were visible in the simulated video. Once again, the participants attempted to identify and mark the targets. This time, however, all experimental elements were integrated into a single window. The craft appeared in the context of the terrain it was navigating, the video was semi-projected onto that terrain, and the participants marked targets directly in the synthetic environment by left-clicking with the mouse where they observed a target.

The independent variable in this study was the virtual camera perspective behavior. The virtual camera began either in chase perspective or north-up perspective (see Figures 4.13 and 4.14). Half-way through the flight, the virtual camera would either transition from one perspective to the other or continue in the same perspective. If the camera transitioned, it followed one of three transition models: instantaneous, smooth, or two-axis smooth. The two-axis smooth condition separated the necessary virtual camera rotations into two components (azimuth and elevation) and gradually changed one at a time while smoothly shifting the virtual camera to the correct location. The smooth transition also gradually changed the virtual camera position and angle as necessary but did so in the shortest single motion possible.

Participants were instructed to remember the flight path. At the end of the flight, the interface perspective smoothly zoomed out to show the entire map and instructed the participants to do their best to trace the path the UAV flew. Once satisfied with the flight path estimation, the participant pressed a button to continue to the next trial. Each flight covered the same distance, had no overlap, and consisted of five straight segments with four turns of either forty-five degrees or ninety degrees. Participants observed both control cases (always chase perspective or always north-up perspective), and three other cases, one from each type of transition, in a randomized order.

We were interested in the effect of different transitions on path recall and target identification. We hypothesized that an instantaneous transition would be most disruptive to path recall because of its disorienting effect. We also assumed that transitions would briefly affect target identification accuracy. We further hypothesized that instantaneous transition would be most disruptive and that the two-axis smooth transition would be the least disruptive (after the control case of no transition). We believed that the smooth transition would reduce the need to reorient by keeping the data in context (showing the relationship between the two perspectives) and that separating axes of rotation would support a gravity-based mental model.

We quickly found that the participants were generally incapable of remembering the automatically executed flight path while focusing on the identification/detection task. Paths seemed almost completely random and participants admitted that they had no clue what the actual flight path was. We tried allowing subjects to use a paper and pencil to help with remembering the flight path. With a paper for taking notes, participants performed better at remembering the shape of the flight path, but had very little sense of scale or location or even the relative lengths of the five flight segments. This indicates that they did not know where the craft had actually flown, but just that it had made certain turns. We attempted to

give a sense of scale and location by showing on the map where the craft started and stopped but participants still could not recreate the flight path with any degree of recognizable accuracy.

Because the recall task was so difficult the data was not very useful. By itself, the target identification task was not very interesting. It was rather easy except during the transition, but the transition only happened once and so briefly that there could only be a very small effect. Moreover, in some flights the craft was flying north when the camera perspective changed. This made the two-axis smooth transition behave the same as the smooth transition (there was no azimuth change to be made). Perhaps the most significant finding from this study was in the subjective data: several participants mentioned that they disliked the instantaneous transition and that it was confusing.

Although the effect of different virtual camera transitions on working memory is interesting, we expect relatively few perspective transitions during a normal flight. Most time should be spent analyzing video with a little attention spent controlling the flight path. One of the main purposes of our research has been to create an interface that a WiSAR volunteer could use to control a UAV to assist with searches. We therefore designed another experiment in which we studied simultaneous path control and target detection.

5.2 Perspective experiment

In this experiment, we explored the effect of virtual camera perspective on a reactive search task using a limited-functionality version of the interface described in Chapter 4. Many of the control options were disabled for the purpose of experimental control. We studied how well an operator with minimal training could perform a search task while operating the interface using four of the most common control perspectives described in Section 4.3.2: chase, north-up, track-up, and full map (see

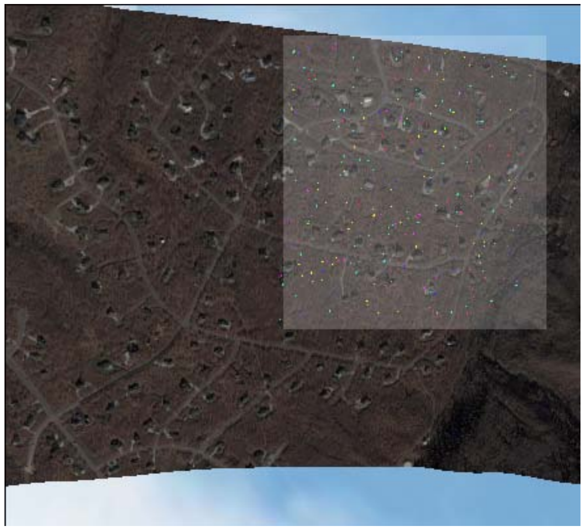


Figure 5.3: Uniform distribution

Figures 4.13 through 4.15). Participants used each of the different perspectives to find targets randomly distributed according to each of three different distributions: uniform, Gaussian, or rectangular path.

5.2.1 Design

One design goal was to test how well an individual without previous experience with the interface could use it to perform a reactive search. We also wished to experiment on the relative usefulness of different perspectives for different types of search. We selected three different probability distributions that we felt suggested different types of searching. Having targets distributed uniformly across a sub-region of the terrain suggests a constraining search to find the distribution area limits and then an exhaustive search of that area. When time is constrained, having targets scattered according to a Gaussian distribution suggests a high-probability, prioritized search pattern. Having targets distributed closely along a constrained path suggests a hasty search.

To provide interactive control for this experiment, the interface connected to Aviones, a moderate-fidelity simulation created by Morgan Quigley that runs the

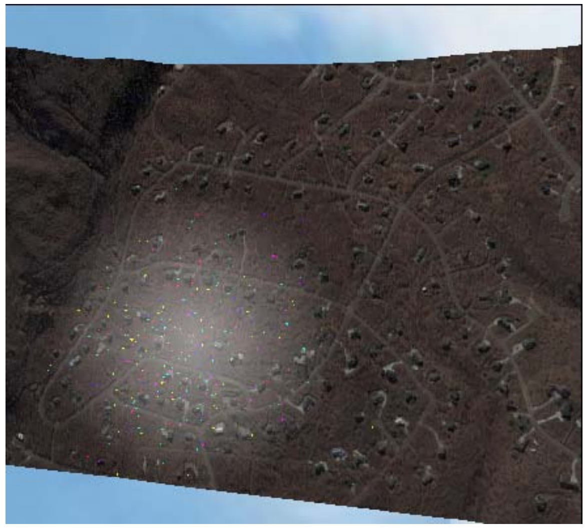


Figure 5.4: Gaussian distribution

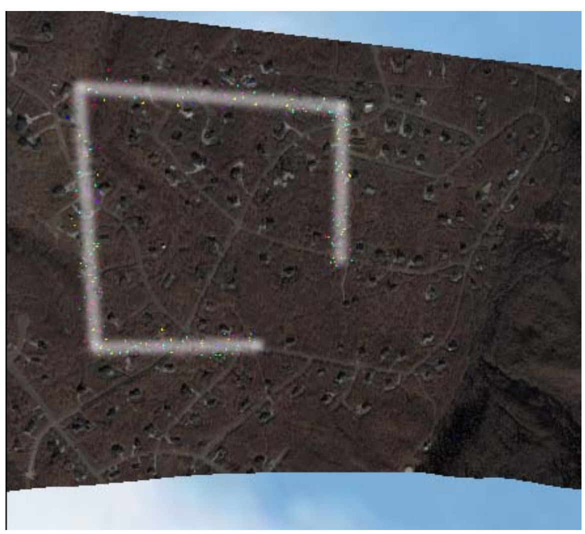


Figure 5.5: Rectangular path distribution

same autopilot code as the physical craft and uses a flight-dynamics physics model to simulate the craft's behavior. The simulator generates imagery as it would be seen by the UAV camera using a synthetic terrain model very similar to that used by the interface (see Section 4.3.1). The simulator accepts commands from the interface and sends back telemetry information and live video.

Test participants were given a sheet of directions introducing them to the interface, instructing on its use, and explaining the experimental task. Subjects participated in twelve experimental trials and four practice trials for a total of sixteen trials. After each experimental trial, participants answered three questions about the relative difficulty of the task and then went on to the next trial. The study ended with a few more general questions about the interface.

Each of the sixteen trials took place in synthetic environments modeled after different locations. The environments were all similar, with a large flat central area and small hills off to the sides. Each participant controlled the craft through all four experimental perspectives. The perspectives were in randomized and counterbalanced order. For each perspective, the participants began with a practice trial to learn how to use the different controls in the perspective. The simulator populated the terrain with only targets (colored spheres), and then the interface gave the participant one minute to practice using the actions available during the experiment: controlling the craft, taking snapshots, and marking targets. When the practice trial was over, the participants performed three experimental trials using the same perspective. Each of the three experimental trials under this perspective used one of the three target distributions, again in randomized, counterbalanced order.

In each experimental trial, both the simulator and the interface began by loading the next terrain model. The simulator then populated the terrain with colored spheres and pyramids randomly scattered according to one of the three different distributions: uniform, Gaussian, or rectangular path (Figures 5.3 through 5.5). Both

Figure 5.6: “Carrot” marker used to guide the UAV

spheres and pyramids followed the same distribution but there were 300 pyramids and 10 spheres. Subjects were instructed to locate and mark the spheres. The pyramids served as distractors (to keep the participant from simply marking any object that stood out from the brownish terrain imagery). Pyramids also indicated the probability distribution so that if there were a large number of pyramids in an area, it was more likely that there was a sphere in the same area. The pyramids fill the role of minor environmental clues such as game trails or vegetation that may not appear in satellite imagery, but give some hint about where a more important clue may or may not be when seen through the live video.

After the trial was setup with the current perspective and target distribution, participants pressed the Enter key to launch the craft. Subjects directed the craft with the mouse using a stick and carrot metaphor. The “carrot” was a distinct marker (Figure 5.6) rendered onto the synthetic terrain that would follow the mouse cursor as long as the Control key was down. When the test subject released the Control key, the marker stayed where it was and the craft continued to fly toward it. When the craft arrived at the marker, it first crossed over the point and then began to circle until the marker was moved. Typically the onboard camera pointed thirty degrees forward from straight down (with respect to the craft), but when the UAV began circling a point, it focused on that point. This same control method was used for all four perspectives.

Marking the spheres was accomplished by using the mouse to left-click on the terrain location where the subject believed the sphere to be. When participants marked a location, a spherical marker stayed in that location. Performing a left-click on an existing mark allowed the subject to drag the mark around, while performing a right-click deleted the mark. Participants also had the option of pressing the space bar to take a snapshot of the video. The snapshot left a still frame of the video at the location the camera was pointing to when the snapshot was taken. Taking snapshots was not necessary for the task, but was a tool participants could use if they chose in order to get a better look at the video of a particular location or to help mark where the craft had been.

After four minutes, the satellite terrain imagery in the interface faded to black and the interface stopped accepting commands. A message appeared indicating that the trial had ended and instructing the participant to answer the relevant survey questions while the next terrain model loaded. The experiment took place on a 19-inch LCD monitor for the primary interface and a five-inch auxiliary LCD monitor that showed the untransformed video (see Figure 4.15). Participants used a regular optical mouse and three keyboard keys (space bar, Control, and Enter) to perform the experiment. Twenty-one naïve human subjects participated in the experiment. Subjects were reimbursed $12 for their time.

5.2.2 Results

In spite of the practice session before using each perspective, subject performance still shows a strong learning effect in all areas. Figure 5.7 shows that true positive marks generally increase while the subject uses a particular control mode and fall slightly when the participant switches to a new perspective. False positive and redundant marks fall fairly consistently over time, rising slightly with the perspective changes. The fact that performance decreases slightly with each perspective change

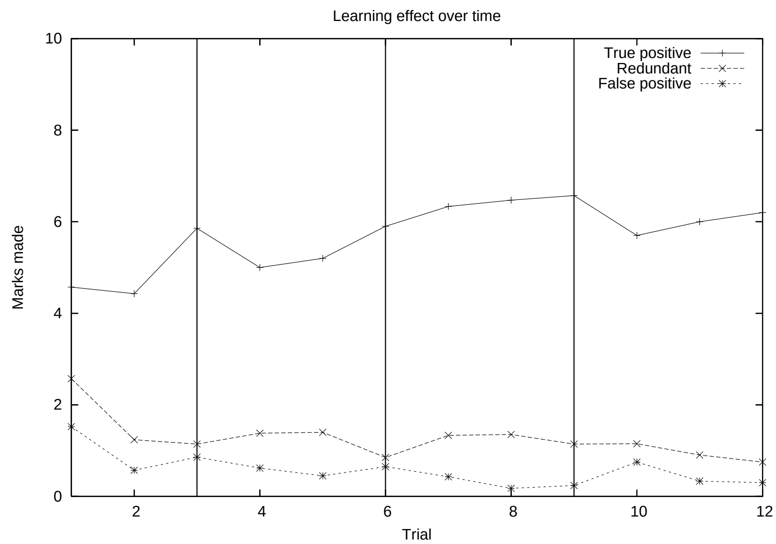
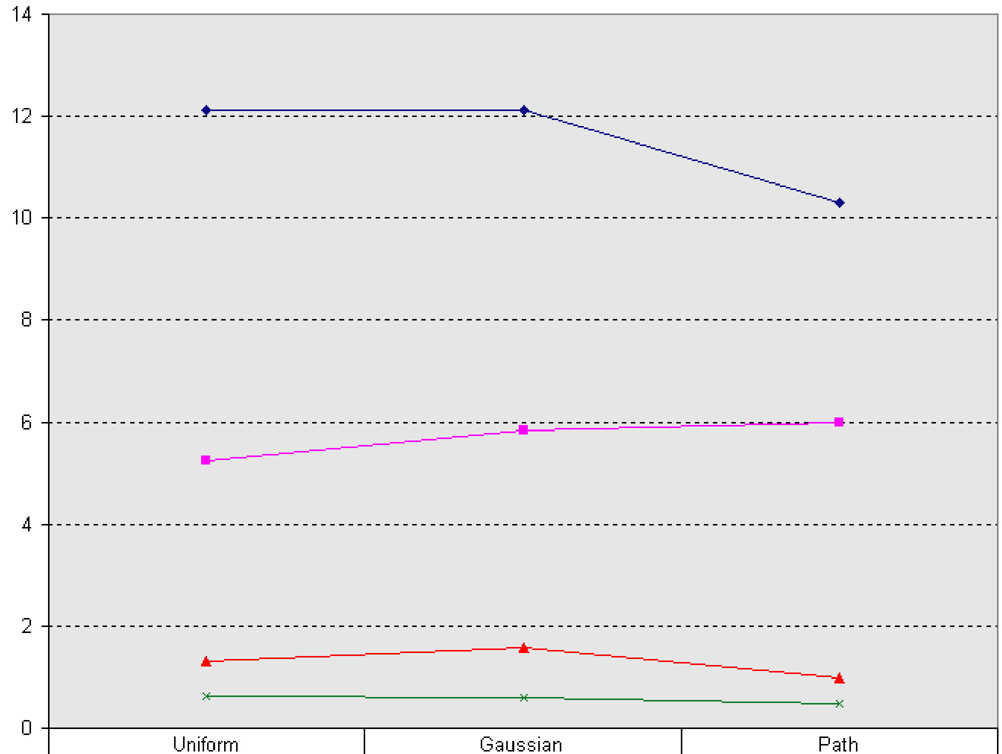


Figure 5.7: Learning Effect

even though everything else remains constant shows that the perspectives are distinct enough from each other to have different strengths, but otherwise, the learning effect is expected and not remarkably interesting.

Other independent variables demonstrate notable and significant effects on performance after statistical analysis (a Tukey-Kramer ANOVA using subjects as a block). Both perspective and distribution significantly affect redundancies, true and false positives, and accuracy of true positives. Figures 5.8 through 5.13 show various performance measures (after Tukey-Kramer adjustment). Figure 5.8 shows a summary of performance according to the three distributions. Figure 5.9 shows a summary of performance data by perspective. Figures 5.10 through 5.13 split the data according to performance metric. Data are grouped by perspective and then distribution.

The data show that the three distributions vary significantly in difficulty. Performance is generally best for the path distribution and worst for the uniform dis-

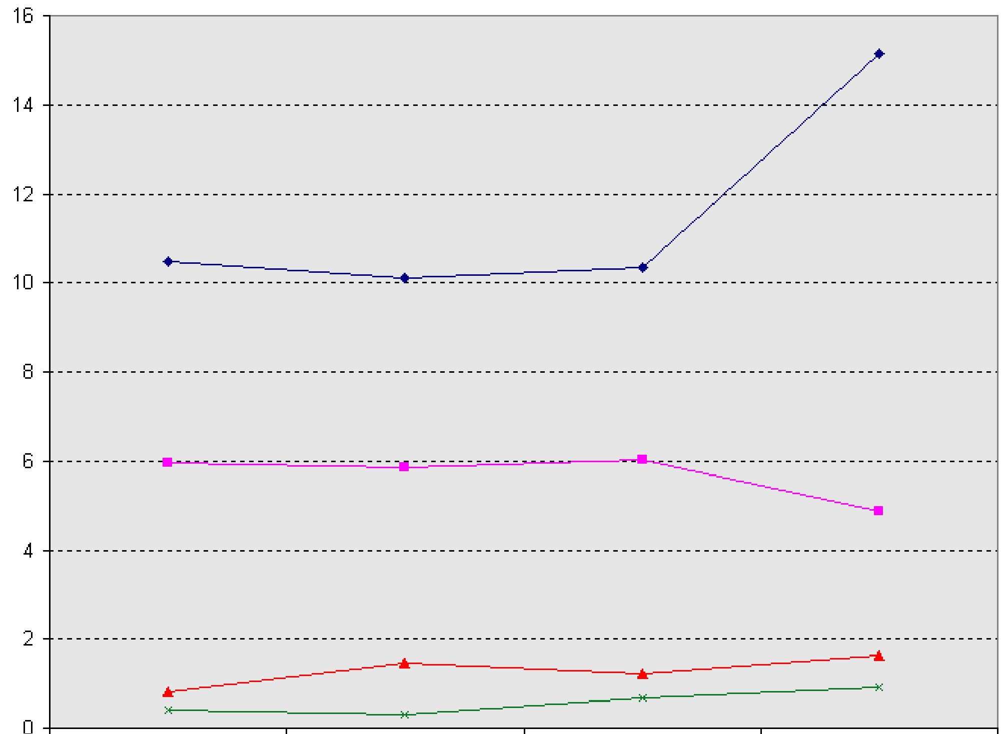


	Uniform	Gaussian	Path
Inaccuracy	12.119	12.12	10.31
True positives	5.238	5.846	5.979
Redundant	1.306	1.578	0.976
False positives	0.638	0.602	0.476

Figure 5.8: Performance means according to distribution

tribution (see Figure 5.8). The uniform distribution demonstrates more redundant marks than the path distribution (p=0.0435) and fewer true positives (p=0.0263).

One reason that path distribution may be easier is that the path distribution suggests an obvious coverage strategy: find and then follow the path. Following the path quickly covers the full probability distribution. Searching the Gaussian distribution from the center outward quickly accumulates probability at the beginning and gradually tapers off with time. Finally, a uniform probability distribution over a rectangular area can be accumulated at a constant but somewhat slow rate. Another implicit advantage of a path distribution is that it is significantly easier to keep track of what part of the distribution has been covered and what has not, leading to less



	Chase	North-up	Track-up	Full-map
Inaccuracy	10.479	10.103	10.332	15.151
True positives	5.973	5.866	6.036	4.876
Redundant	0.806	1.469	1.235	1.638
False positives	0.399	0.301	0.679	0.908

Figure 5.9: Performance means according to perspective

redundant coverage, which means fewer redundant marks. Lower redundant coverage means greater novel coverage and consequently more true positives are found.

This suggests that using a reactive control model such as the stick and carrot metaphor may be best suited for a hasty style of search. It may be more appropriate to use automatically generated search patterns for high-probability or exhaustive searches, with less direct control or intervention. Reactive control may still be effective for a constraining search. Participants seldom attempted to constrain the area but rather tended to fly criss-cross patterns over both the uniform and Gaussian areas, turning around when they stopped seeing pyramids.

The different perspectives also demonstrate a significant effect, although it is not as strong as we had expected. The primary observation is that the full-map

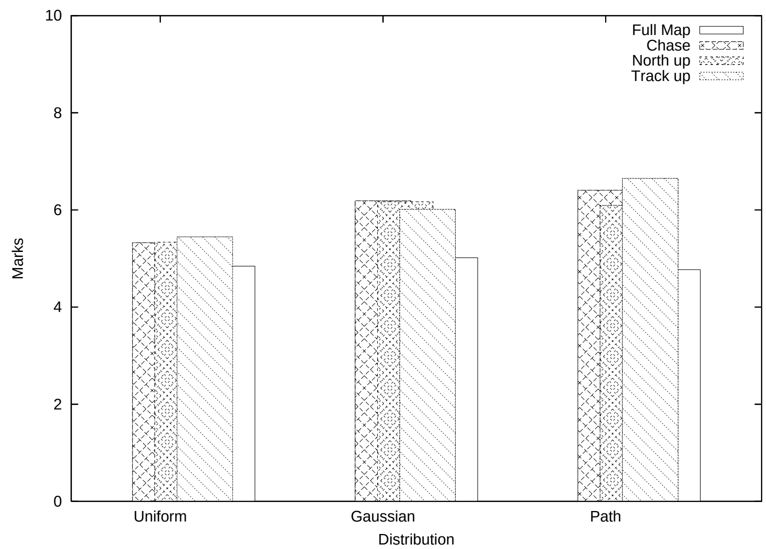


Figure 5.10: True positive marks

perspective is significantly worse ($p < 0.05$) in all respects except redundant marks. All perspectives except full-map show comparable levels of true positive marks (Figure 5.10). In the subjective data, participants rank full-map as more difficult than all other perspectives ($p < 0.0005$). Chase was also ranked as easier than track-up (p=0.0633) and insignificantly (p=.5176) easier than north-up. Overall, subjects performed comparably well using the chase, north-up, and track-up perspective. This is notable because other studies have found improved performance and operator preference using a track-up perspective [28, 65]. This may be because in other studies, they used a traditional control method where commands are given with respect to the craft (e.g., turn right or left). A track-up perspective helps the operator avoid confusing his or her own left with the craft's left. The carrot and stick control metaphor, on the other hand, is terrain-centric; so a moving terrain model can make control more difficult.

Keeping the terrain model completely stationary requires a perspective sufficiently distant to show the entire operating area at once. The full-map perspective

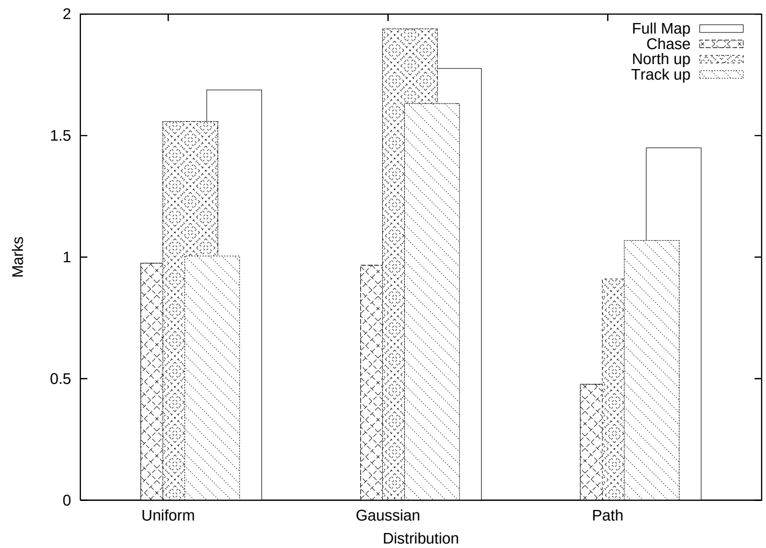


Figure 5.11: Redundant marks

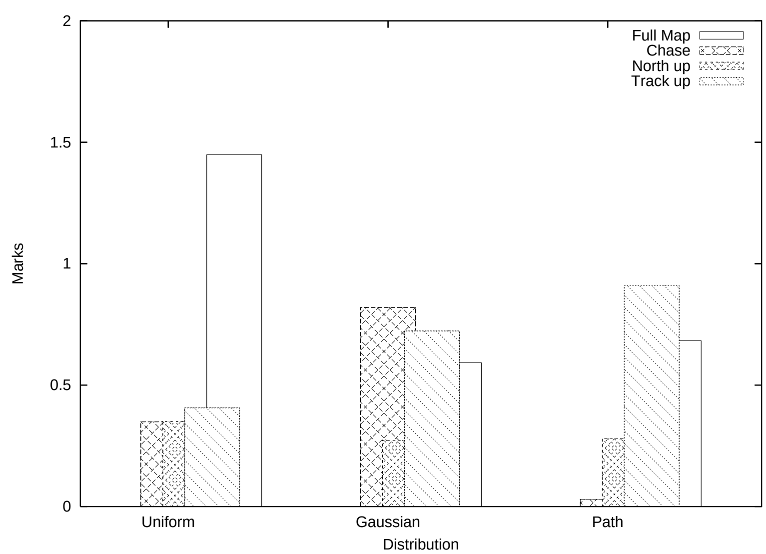


Figure 5.12: False positive marks

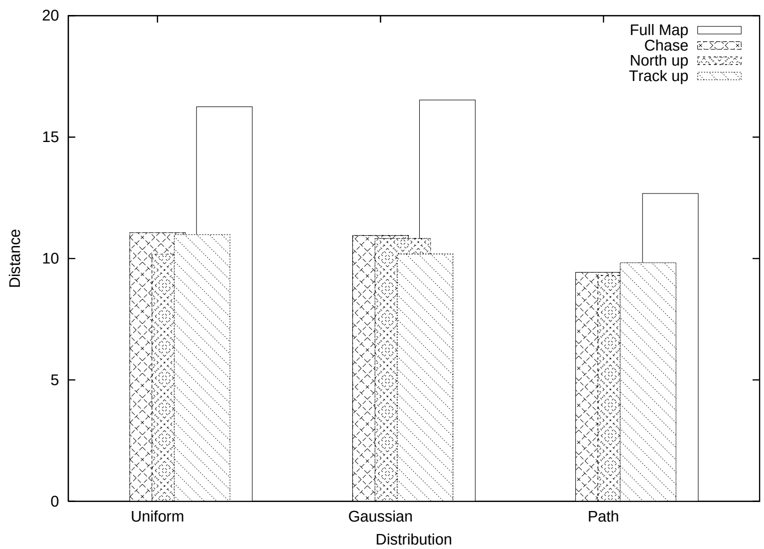


Figure 5.13: Mean distance from true positive actual location

does this and most closely imitates the status quo in camera-equipped UAV interfaces. In the full-map perspective, the video footprint is still visible, but without sufficient pixel resolution to distinguish details. Consequently, participants had to rely on the separate monitor with the raw video in order to detect targets. Many participants commented that they only used the raw video monitor for the full-map perspective and that they disliked it. Participants had to direct the craft on the interface screen and then turn their attention to the video screen to watch for targets. Upon detecting a target, they returned their attention to the interface screen and searched for the video footprint in order to mark the object on the terrain. Marking accurately required mental rotations to correlate the video with the terrain.

Several participants used the snapshot feature for the full-map perspective trials. Participants could concentrate on the raw video monitor with one hand on the mouse and the other on the keyboard. When a target appeared in the raw video monitor, participants took a snapshot by pressing the space bar. They then switched their attention briefly to the primary monitor, found the snapshot, and placed the

sphere mark. Participants who used this strategy generally did better with the full-map perspective than those who did not, but still worse than with other perspectives. This supports our claim that traditional UAV interfaces may not be appropriate for WiSAR.

5.3 Field trials

Experiments in simulation demonstrate many useful principles, but in the field we find many effects and problems that may not show up in simulation. A series of experimental field trials, some more successful than others, have taught us several things about UAV-assisted wilderness search. In these field trials, an individual experienced with WiSAR designed and setup a scenario somewhat typical of the kind faced by first responders. At an appointed time, the researchers involved in this project met at the field site: public land in a remote area where other people and property would not be endangered by a possible malfunction. After equipment was setup and tested, the individual responsible for designing the scenario described the situation as though it were a call recently received at the sheriff's office. Ron Zeeman, an experienced WiSAR volunteer, would then act as incident commander for the trial.

The incident commander and UAV operator would plan out a course of action and then deploy the craft. The UAV operator was always a student with some, but typically not extensive, experience controlling the UAV. The operator controlled the craft through one of the various interfaces under development through this and other UAV projects. Once the craft was deployed, several people, including the operator, monitored the video in search of details or colors revealing possible information about the "missing person's" location. The "missing person" was typically a pair of pants and a t-shirt lying somewhere on the terrain, occasionally accompanied by a bicycle or hiker's backpack (see Figure 5.14). In the area, there might also be a discarded hat

Figure 5.14: A field trial "victim"

or jacket or bicycle tracks which would indicate the missing person's passage through the location.

As the UAV covered different areas, the incident commander would ask to see some areas again or closer. Other times the incident commander might change the plan and decide to look somewhere else. Sometimes the operator and team managed to locate the "missing person" and sometimes things went badly and we had to quit early.

Following each trial, both successful and unsuccessful, the entire group met to debrief the experience. Each researcher independently filled out a subjective survey, rating the technology and discussing the strengths and weakness of the technology. The entire group then discussed what had happened, why it happened, and how it might be improved. The data gathered from these experiments and discussions indicates that there are several different possible models for incorporating UAV-enabled teams into a WiSAR framework as discussed in Section 3.3.5. We found an emphatic though obvious need for a robust platform. We also recognized the need for a high level of neglect tolerance in the system, enhanced video presentation, robust communication links, and clearly organized procedures and responsibilities.

5.3.1 Neglect tolerance

A single UAV operator working in the field is subjected to a number of distractions from controlling the craft, not the least of which are monitoring the video stream and interacting with the rest of the search team. In practice, it may not be feasible to eliminate these distractions. Instead, the system must be made tolerant to such distractions. The system must have a moderate level of neglect tolerance. After insuring the safety of the operator and other search team members, the first priority which must be made neglect tolerant is the task of keeping the craft in flight. The autopilot, when functioning correctly, takes care of this job reasonably well in suitable weather and terrain conditions. Once the craft is airborne and searching near terrain, some height above ground maintenance is imperative.

The intensity of the situation often draws the operator's attention away from the task of monitoring the safety of the craft. In particular, if the same individual is directing the craft and monitoring the video stream, the operator's attention may be focused more on what the video shows than on potential threats to the craft. During one field trial the operator was interested in getting a better look at a particular location and so set up a coverage pattern by placing waypoints to the north and south of the location. To get more detail, the operator decreased the craft's altitude. The commanded altitude was safe for the endpoints of the coverage pattern, but there was a ridge in between. The operator became so engrossed in watching the video that he failed to notice the ground coming up to meet the UAV and did not hear when others tried to alert him to the danger. The craft finally planted itself on the side of the ridge and brought an early end to the field trial.

Flight into a tree or mountainside caused by flying too low brings the search to a rapid halt. On the other hand, high-altitude flight can cause problems by limiting the detail viewable by the fixed-focal length camera. At an altitude of h meters with a view-angle of θ and a camera resolution of d, a target that presents a round

profile with a radius of r meters will present $area \approx \pi * \left(\frac{r*d}{2*h*tan\left(\frac{\theta}{2}\right)}\right)^2$ pixels if full-resolution video is presented to the operator. For example with our standard setup of a 40 degree wide camera with capturing 640 pixel-width video, when flying 100 meters above ground, we would expect a round target with a half-meter radius to be represented by approximately 60 pixels or $\left(\frac{1}{5120}\right)^{th}$ of a full (640x480) video frame [27]. Probability of detecting a visual target is dependent on a great many things, but it decreases quickly with size. Consequently, without an adjustable zoom camera, flying too high can make the video signal almost useless.

We implemented an open-loop attempt at maintaining height above ground. The algorithm is simple: the UAV sends its GPS coordinates and altitude to the ground station. The ground station looks up the terrain altitude at that location using the digital elevation map that is part of the synthetic environment. The interface computes the current height above ground by comparing craft altitude to the terrain altitude. If the height above ground of the craft is more than a couple meters different from the desired height above ground, the ground station automatically sends a new desired altitude to the craft as necessary to correct the discrepancy (e.g., go higher if the craft is too low). This naïve approach performs very well over relatively gradual changes in terrain and contributed to the success of two subsequent field trials.

Although this simple height-above-ground maintenance is a vast improvement over nothing at all, it suffers from several limitations. First, because the terrain information is on the ground station and not onboard the autopilot, if communications are spotty, the craft may not receive important altitude corrections or may go into a problematic failsafe mode. During one field trial, the craft was climbing over a ridge when it seems to have temporarily lost communications with the ground station. It engaged a fail-safe mode that tells it to maintain a height of 100 meters above launch altitude and fly back to launch point. Unfortunately, in these circumstances, 100 meters above launch altitude happened to be below ground height. The craft

descended while turning toward launch point and promptly ran into the only large boulder on an otherwise sandy mountain.

Another problem that factored into this crash was that the open-loop height-above-ground maintenance does not account for maximum climb rate or look ahead at all. The slope of the mountain increased faster than the craft was climbing, consequently bringing the craft closer to the terrain than the operator intended and increasing the severity of the loss of altitude incurred when the craft entered fail-safe mode.

5.3.2 Persistent, enhanced, terrain referenced imagery

In one of our first field trials, we went out with optimistic expectations of quickly locating the target and being back home after just a couple hours. We were disappointed. After a lengthy and frustrating series of mechanical and electrical failures, the UAV was flying and we began to search around the missing person's point last seen. During this trial, the ground station used a traditional interface model with map-based control on one display and a separate screen for monitoring the video. The operator had the craft circle various areas around the point of interest while the rest of the team crowded around the monitor and argued about the video. With the craft flying circles in a stiff wind, the video shook so much that it was very difficult to discern anything. Something looked like it might be a person. That was good enough for the overanxious search team. A bunch of people took off to go inspect the general area where the craft was. Meanwhile a few people stayed back to try to get a better view and a better estimate of the location.

The group watching the video could not find the object of interest again, nor could they decide if what they were looking at was the same thing as before. When the field team arrived near the area and asked for further directions, the base team could not give any. In the end, whatever it was that had shown up in the video was

not the missing person or even remotely close. However the experience highlighted several difficulties associated with using the traditional UAV interface setup to search.

High frequency jitter introduced by the instability of the craft can make it difficult to focus on anything of interest. When the camera is focused on a small enough area to make out significant detail, a small object is only in view for a brief moment making it hard to localize. As the craft circles a point, it is very difficult to avoid being disoriented because there is no easy way to follow how much the craft has turned. Finally, it is quite difficult to pinpoint the exact location shown by the video because it requires integrating the craft GPS location, altitude, heading, pitch, roll, camera angles, and terrain information. This level of mental gymnastics is very difficult for a human to do in real time, but is trivial for a machine.

With training, humans can overcome this sort of difficulty to some degree [42]. However, technological improvements can also make the task easier and less error prone. Damon Gerhardt used some basic computer vision techniques to remove high-frequency jitter from the video [18]. Incorporating this into our video display made a big difference in clarity. Damon also developed a way to stitch several seconds of video into a small mosaic that increases the time available for inspecting imagery and keeps imagery aligned with a constant direction even if the craft turns. He found that this can make a huge difference in a detection task [18]. Determining and demonstrating where on a map the UAV camera is pointing is a simple problem for a computer. Later field trials benefited significantly from these technologies.

5.3.3 Reliable communication lines

Successfully operating the craft requires reliable communication links. The command/telemetry link is essential. Without it, the ground control station has no way of sending commands or knowing the state of the craft. The command/telemetry link is accomplished over a radio modem that has limited range and typically requires line-

of-sight. When the telemetry link fails, the UAV typically turns around and flies back toward launch point. This may cause problems, as mentioned in Section 5.3.1, but the operator has no way to avoid them while communications are down. Problems with comms have been a significant source of trouble in field trials.

Failed communication links between separate field teams have also cause problems. During one trial, the field team left base camp in order to be in position before the base team deployed the craft. Both base and field teams had radios, but a mountain disrupted line of sight between the teams. After being unable to contact the field team by radio or cell phone for several minutes, the base team decided to deploy the craft and begin executing the search plan. When a failure in the autonomy crashed the UAV on a mountain, the base members left to retrieve it, leaving base camp unattended. This resulted in a bad situation where team members could not communicate, did not know where each other was, and had no way to find each other. Having reliable communications and a protocol for reestablishing them can improve team efficiency.

5.3.4 Organized plan

No doubt the trained volunteers on Search and Rescue teams realized this long ago, but for mission success it is imperative to be organized. Without organization, those conducting the search may expose themselves to unacceptably large risks. In the case described above, with the teams separated from each other and no plan for reconnecting, if an individual had actually gotten lost, there could have been a real wilderness search and rescue situation complete with all the dangers to the searchers and the missing person. A plan for when to abandon the UAV and how to behave in case of various eventualities can protect the entire WiSAR team from unnecessary risk.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Small, camera-equipped UAVs have the potential to offer substantial support in the WiSAR domain because of their ability to rapidly acquire imagery of wilderness areas. Small UAVs can be rapidly deployed at less expense than manned aircraft and without endangering a human pilot. The research described in this thesis is incremental work toward making this a reality through formal analysis to determine domain specific requirements and constraints, followed by human-centered design to meet these requirements in a reasonable manner. The design has been partially validated through controlled experimentation and partially-structured field trials, which also demonstrate some general principles of UAV control systems.

Through formal analysis, we have developed a model for how WiSAR is currently accomplished and how it might be supported with small, camera-equipped UAVs. The analysis shows that the key activity in WiSAR is gathering information that directly or indirectly leads to evidence of the missing person's location. A camera-equipped UAV can serve as a tool for acquiring information from wilderness areas but it also introduces additional tasks of deploying, monitoring, controlling, and retrieving the UAV. Furthermore, the system must be portable, neglect tolerant, and simple to use, while providing useful imagery.

A portable craft and ground-station hardware that meet WiSAR constraints already exist. Through human-centered design, we are developing software to support WiSAR needs while accounting for human abilities and limitations. Limits on human sensory and cognitive processing imply that we must be careful with how we present the information. The presence of distractions and human error imply that we must have automatic routines to minimize consequences when the human neglects the system or makes a mistake and to simplify flight details that do not directly concern the search task. Automation on the craft and ground station can help a WiSAR volunteer to deploy the UAV, keep the UAV in the air, systematically search an area, and finally retrieve the UAV. Ecological presentation of terrain, craft, and video support situation awareness and provide an intuitive model for reactively controlling the flight path.

Our experimental validation of the interface has shown that a traditional UAV interface model with separate windows for map control and raw video is not appropriate for WiSAR. We found that test participants were less effective when searching with a full-map perspective and separate video source than with an integrated display that showed both terrain and video. Perhaps most importantly, we observed that participants were capable of controlling the simulated UAV to perform a search after less than ten minutes of instruction. Up to this point, our validation efforts have only included a small portion of the interface design. There are many experiments to be done in the future and several more refinements to be made to the interface software, but we have shown that it is possible to create a system that allows a single operator to use a camera-equipped UAV to perform a search task for Wilderness Search and Rescue with only minimal task-specific training.

6.2 Future work

We have explored several interesting questions, but many more can be studied using the interface framework we have developed. This thesis represents incremental progression toward the knowledge necessary to build a fully-functional UAV system capable of enabling WiSAR volunteers to use camera-equipped UAVs for search. Several steps remain to be taken before the research presented in this thesis can be deployed to support WiSAR. Some necessary technologies already exist and must only be integrated into a single interface. Other technologies still require significant exploration, development, and refinement.

6.2.1 Multi-agent interface extension

The design of the system is such that updating to a multi-agent application would not be extremely difficult. Modular design makes it so that one would only need to make a few minor changes to underlying code and then instantiate multiple instances of the object representing the craft. However, the logistics of a human actually interacting with multiple crafts need to be studied. In order for a human operator to manage multiple instances of a craft, the system would require sufficient automation to provide the neglect tolerance necessary to allow an operator to make effective use of the different crafts [20]. This suggests the need for more advanced automation that allows the operator to give more abstract, long-term commands. It would also require some mechanism for controlling the temporal demands of inspecting video because it is impractical to expect anyone to pay attention to multiple frames of video simultaneously. Real-time mosaicking of multiple video sources may eventually be able to compress a significant length of time and several different videos into a single image that can be inspected as time allows. Flight automation will allow the operator to designate high-priority areas of interest and then monitor the progress of

several craft as they negotiate how to cover the areas and then return the requested imagery.

6.2.2 Integration with mosaic

Damon Gerhardt and Dr. Bryan Morse, who developed the video stabilization algorithm currently used in the interface software described here, have also developed and studied the ability to mosaic several frames of video. This changes the search task from nearly instantaneous ($\frac{1}{20}$ second image persistence) to a few seconds. The operator monitoring the video can now blink without missing an artifact in the video stream. Just a few seconds of persistence make a tremendous difference. In their study, Morse and Gerhardt found a 43 percent higher correct-detection rate when using a short term mosaic with only a small corresponding increase in false positives [18]. We expect that incorporating this technology into the interface will offer similar improvements to detection in a search task and may provide other benefits as well.

6.2.3 Full 3D interaction model

As an exploratory interface, many desirable features have not been fully implemented. Many others still require testing. Because the interface uses a synthetic environment to present information about the UAV within the context of its environment, many of the presentation elements are displayed using 3D rendering techniques. Interacting with 3D icons is different from interacting with 2D icons. Mouse actions are reported to the software as an ordered pair that gives the location of the pointer on the 2D screen. It is trivial to test a 2D rectangular icon to see if it contains the 2D point that is the mouse cursor. However, the addition of a third dimension not only introduces an ambiguous axis, but with larger space, there tend to be more objects to check. Several techniques exist for selecting 3D objects. The current software uses ray-picking to

recognize what part of the terrain-model is under the mouse cursor. This, or some other method, can be used to select and manipulate waypoints, the UAV, and other iconic objects in the interface. However, a way must be devised to disambiguate axes when dragging in 3D or attempting to click on an icon that is occluded by another icon.

6.2.4 Playback functionality

It would be advantageous to be able to pause, rewind, and fast-forward the progress of the flight (with video up to the present, and the predicted state thereafter). With the proper setup it would be possible to play multiple portions of the flight simultaneously and thus monitor the current progress of the flight while also replaying another portion of the flight. Some evidence shows that the ability to replay may be undesirable in some circumstances because it causes people to miss the present [52]. However, as automation improves to increase neglect time of the system, the operator will have more leeway to slowly scrutinize portions of the flight that merit careful inspection, and then quickly scan through portions that clearly contain little of interest.

6.2.5 Sophisticated 3D path planning

The planning used onboard the craft is fairly simplistic. Even a small amount of planning makes a big difference in the workload on the operator. As the automation becomes more powerful and more reliable, the craft will become more useful. Researchers in the HCMI lab are working on statistical methods for estimating the utility of searching sub-regions of an incident site. We are also developing heuristic approaches for optimizing flight time given an estimate of the utility for searching different regions of an area. This sophisticated path planning will likely lead to more effective use of the UAV as a search resource.

6.2.6 Airspace integration / Meeting FAA regulations

The Federal Aviation Administration is currently attempting to develop appropriate policies for regulating the use of unmanned aircraft. One difficulty is that UAVs vary drastically. Some UAVs are the same size as commercial aircraft. Some are smaller than many birds. Because the field has recently begun to rapidly expand as a field of active research, things are in flux and there is a great demand for the technology. However, the FAA wishes to avoid injury to life or property through the new technology and is developing strict regulations for controlling any unmanned aircraft [2]. When a final system is implemented for actual WiSAR use, it will be important that it comply with legal regulations and avoid endangering other aircraft as well as life and property on the ground.

6.2.7 Integration with other WiSAR technology

Section 3.3.5 discussed using the UAV system as another technical search specialist similar to the man-tracking specialist or canine specialist. However, information from the UAV could also be combined into a data integration system. It is feasible within the next several years to develop a system that not only tracks and organizes the progression of multiple UAVs, but also records the path and findings of other search teams. Ground-based search teams already carry beacons that transmit their progress through a search. The system would need a way for incident command to annotate the map with information and dynamically update probability maps with the passage of time.

UAV technology has tremendous potential to help save the lives of individuals who get lost in the wilderness. We hope that our work will help make this happen as well as contribute to the general knowledge of human-robot interaction.

Bibliography

[1] J. Adams, J. Cooper, M. Goodrich, C. Humphrey, M. Quigley, B. Buss, and B. S. Morse. Camera-equipped mini UAVs for wilderness search support: Task analysis and lessons from field trials. Technical report, BYU and Vanderbilt, 2007. URL: http://tanglefoot.cs.byu.edu/.

[2] Federal Aviation Administration. Unmanned aircraft systems operations in the U.S. national airspace system—interim operational approval guidance. Memorandum, September 2005. AFS-400 UAS Policy 05-01.

[3] A. L. Alexander, C. D. Wickens, and D. H. Merwin. Perspective and coplanar cockpit displays of traffic information: Implications for maneuver choice, flight safety, and mental workload. *International Journal of Aviation Psychology*, 15(1):1–21, 2005.

[4] L. Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, 1983.

[5] D. B. Barber, S. R. Griffiths, T. W. McLain, and R. W. Beard. Autonomous landing of miniature aerial vehicles. In *AIAA Infotech@Aerospace*, number AIAA-2005-6949, 2005.

[6] Alan F. Blackwell. The reification of metaphor as a design tool. *ACM Trans. Comput.-Hum. Interact.*, 13(4):490–530, 2006.

[7] J. Blinn. Where am I? What am I looking at? *IEEE Computer Graphics and Applications*, 8(4):76–81, 1988.

[8] G. L. Calhoun, M. H. Draper, J. T. Nelson, and H. A. Ruff. Advanced display concepts for UAV sensor operations: Landmark cues and picture-in-picture. In *Proceedings of the Human Factors and Ergonomics Society 50th annual meeting*, pages 121–125, 2006.

[9] CERI. Manning the unmanned. In *2005 Human Factors of UAVs Workshop*, Mesa, Arizona, May 2005. Cognitive Engineering Research Instutitue.

[10] CERI. Manning the unmanned. In *2006 Human Factors of UAVs Workshop*, Mesa, Arizona, May 2006. Cognitive Engineering Research Instutitue.

[11] R. Christiansen. Design of an autopilot for small unmanned aerial vehicles. Master's thesis, Brigham Young University, Provo, Utah, USA, 2004.

[12] Utah County. Sheriff's office search and rescue team, August 2005. Web Resource. URL: http://www.co.utah.ut.us/Dept/Sheriff/volunteer.asp.

[13] M. L. Cummings. *Designing decision support systems for revolutionary command and control domains*. Doctoral dissertation, University of Virginia, 2003.

[14] J. L. Drury, J. Richer, N. Rackliffe, and M. A. Goodrich. Comparing situation awareness for two unmanned aerial vehicle human interface approaches. In *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR)*, Gaithersburg, Maryland, USA, August 2006.

[15] M. R. Endsley. A survey of situation awareness requirements in air-to-air combat fighters. *International Journal of Aviation Psychology*, 3(2):157–168, 1993.

[16] M. R. Endsley, B. Bolté, and D. G. Jones. *Designing for Situation Awareness: An Approach to User-Centered Design*. Taylor and Frances, New York, 2003.

[17] D. L. Fisher, E. J. Yungkurth, and S. M. Moss. Optimal menu hierarchy design: Syntax and semantics. *Human Factors*, 32(6):665–683, 1990.

[18] D. D. Gerhardt. Feature-based unmanned air vehicle video euclidean stabilization with local mosaics. Master's thesis, Brigham Young University, Provo, Utah, USA, 2007.

[19] M. Goodrich, B. Morse, M. Quigley, D. Gerhardt, J. Cooper, J. Adams, and C. Humphrey. Augmenting wilderness search and rescue using a camera-equipped mini-UAV. *Journal of Field Robotics*, 2007. In review.

[20] M. A. Goodrich and D. R. Olsen. Seven principles of efficient interaction. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pages 3943–3948, 2003.

[21] M. A. Goodrich, M. R. Quigley, and K. A. Cosenzo. Task switching and multi-robot teams. In *Proceedings of the 2005 NRL Workshop on Multi-Robot Systems*, 2005.

[22] M. A. Goodrich, R. J. Rupper, and C. W. Nielsen. Perceiving head, shoulders, eyes and toes in augmented virtuality interfaces for mobile robots. In *IEEE International Workshop on Robot and Human Interactive Communication, 2005*, pages 574–579, Nashville, TN, 2005.

[23] C. Gutwin, J. Dyck, and C. Fedak. The effects of dynamic transparency on targeting performance. In *Graphics Interface*, Halifax, Nova Scotia, 2003. CIPS, Canadian Human-Computer Communication Society, A K Peters.

[24] K. J. Hanna, H. S. Sawhney, R. Kumar, Y. Guo, and S. Samarasekara. Annotation of video by alignment to reference imagery. In *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms*. Springer, 2000.

[25] B. L. Harrison, G. Kurtenbach, and K. J. Vicente. An experimental evaluation of transparent user interface tools and information content. In *UIST '95: Proceedings of the 8th annual ACM symposium on User interface and software technology*, pages 81–90, New York, NY, USA, 1995. ACM Press.

[26] R. M. Howard and I. Kaminer. Survey of unmanned air vehicles. In *Proceedings of the American Control Conference*, volume 5, 1995.

[27] J. A. Jackson. Panoramic video for efficient ground surveillance from small unmanned air vehicles. Master's thesis, Brigham Young University, Provo, Utah, USA, 2007.

[28] C. Jansen, S. de Vries, and M. Duistermaat. Optimizing the presentation of UAV images in an attack helicopter cockpit. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting—2006*, pages 131–135, San Francisco, CA, USA, 2006.

[29] R.S. Jensen. Prediction and quickening in perspective flight displays for curved landing approaches. *Human Factors*, 23:355–363, 1981.

[30] M. St. John and M. B. Cowen. Use of perspective view displays for operational tasks. Technical report, Pacific Sciences and Engineering Group Inc, San Diego CA, 1999.

[31] J. S. Knutzon. *Managing multiple unmanned aerial vehicles from a 3D virtual environment*. Doctoral dissertation, Iowa State University, 2006.

[32] B. O. Koopman. *Search and Screening: General Principles with Historical Applications*. Pergamon Press, 1980. This book has been reprinted in its entirety in 1999 by the Military Operations Research Society, Inc.

[33] C. A. Miller, H. B. Funk, M. Dorneich, and S. D. Whitlow. A playbook interface for mixed initiative control of multiple unmanned vehicle teams. In *Proceedings of the 21st Digital Avionics Systems Conference*, volume 2, pages 7E4–1 – 7E4–13, Irvine, California, November 2002.

[34] M. Mulder. An information-centered analysis of the tunnel-in-the-sky display, part two: Curved tunnel trajectories. *International Journal of Aviation Psychology*, 13(2):131–151, 2003.

[35] C. W. Nielsen. *Using Augmented Virtuality to Improve Human-robot Interactions*. Doctoral dissertation, Brigham Young University, Provo, Utah, USA, 2006.

[36] R. Parasuraman, T. B. Sheridan, and C. D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, 30(3):286–297, May 2000.

[37] M. Plumlee and C. Ware. An evaluation of methods for linking 3D views. In *Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 193–201, New York, NY, USA, 2003. ACM Press.

[38] M. Plumlee and C. Ware. Integrating multiple 3D views through frame-of-reference interaction. In *International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pages 34–43, July 2003.

[39] M. D. Plumlee and C. Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Trans. Comput.-Hum. Interact.*, 13(2):179–209, 2006.

[40] M. Quigley, M. A. Goodrich, and R. W. Beard. Semi-autonomous human-UAV interfaces for fixed-wing mini-UAVs. In *Proceedings of the International Conference on Intelligent Robots and Systems (IROS), 2004*, Sendai, Japan, Sept-Oct 2004.

[41] M. Quigley, M. A. Goodrich, S. Griffiths, A. Eldredge, and R. W. Beard. Target acquisition, localization, and surveillance using a fixed-wing, mini-UAV and gim-

baled camera. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005.

[42] J. Rasmussen. Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, 13(3):257–266, May/June 1983.

[43] J. A. Ratches, C. P. Walters, R. G. Buser, and B. D. Guenther. Aided and automated target recognition based upon sensory inputs from image forming systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):1004–1019, September 1997.

[44] J. Redding, T. W. McLain, R. W. Beard, and C. Taylor. Vision-based target localization from a fixed-wing miniature air vehicle. In *Proceedings of the American Control Conference*, Minneapolis, Minnesota, USA, June 2006.

[45] R. W. Ricks. An ecological display for robot teleoperation. Master's thesis, Brigham Young University, Provo, Utah, USA, 2004.

[46] H. Ruck. Human effectiveness directorate: Overview. In *Human Factors of UAVs Workshop*, Mesa, Arizona, May 2005.

[47] A. D. Saja. The cognitive model: an approach to designing the human-computer interface. *ACM SIGCHI Bulletin*, 16(3):36–40, 1985.

[48] N. B. Sarter and D. D. Woods. How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37(1):5–19, 1995.

[49] T. J. Setnicka. *Wilderness Search and Rescue*. Appalachian Mountain Club, 1980.

[50] T. B. Sheridan. *Humans and Automation: System Design and Research Issues*. John Wiley and Sons, Santa Monica, CA, 2002.

[51] T. B. Sheridan and W. L. Verplank. Human and computer control of undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

[52] H. S. Smallman and M. St. John. CHEX (Change History Explicit): New HCI concepts for change awareness. In *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*, pages 528–532, Santa Monica, CA, USA, 2003.

[53] H. S. Smallman and M. St. John. Naive realism: Misplaced faith in realistic displays. *Ergonomics in Design: Magazine of Human Factors Applications*, 13(3):6–13, 2005.

[54] H. S. Smallman, M. St. John, H. M. Oonk, and M. B. Cowen. Information availability in 2D and 3D displays. *IEEE Comput. Graph. Appl.*, 21(5):51–57, 2001.

[55] L. G. Strater and M. R. Endsley. SAGAT: a situation awareness measurement tool for commercial airline pilots. In *Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium*, Savannah, GA, 2000. SA Technologies, Inc.

[56] K. S. Tao, G. K. Tharp, W. Zhang, and A. T. Tai. A multi-agent operator interface for unmanned aerial vehicles. In *Proceedings 18th Digital Avionics Systems Conference*, 1999.

[57] J. W. R. Taylor and K. Munson. *Jane's Pocket Book of Remotely Piloted Vehicles: Robot Aircraft Today*. Collier Books, New York, 1977.

[58] K. Vicente. *Cognitive Work Analysis: Toward Safe, Productive and Healthy Computer-Based Work*. Lawrence Erlbaum Associates, 1999.

[59] K. J. Vicente and J. Rasmussen. Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, man, and Cybernetics*, 22(4):589–604, July/August 1992.

[60] W. Wagner. *Lightning Bugs and other Reconnaissance Drones*. Armed Forces Journal International, Fallbrook, CA, 1982.

[61] C. Wickens. Aviation. *Handbook of Applied Cognition*, pages 361–389, 2007.

[62] C. D. Wickens and C. M. Carswell. The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, 37(3), 1995.

[63] C. D. Wickens and J. G. Hollands. *Engineering Psychology and Human Performance*. Prentice-Hall, New Jersey, third edition, 2000.

[64] C. D. Wickens, O. Olmos, A. Chudy, and C. Davenport. Aviation display support for situation awareness. Technical report, University of Illinois Aviation Research Lab Technical Report, 1997.

[65] C. D. Wickens and T. T. Prevett. Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied*, 1(2):110 – 135, 1995.

[66] K. W. Williams. A summary of unmanned aircraft accident/incident data: Human factors implications. Technical report, FAA Civil Aerospace Medical Institute, 2004. URL: http://www.hf.faa.gov/docs/508/docs/uavFY04Mishaprpt.pdf.

[67] D. D. Woods and S. W. A. Dekker. Anticipating the effects of technological change: A new era of dynamics for human factors. *Theoretical Issues in Ergonomic Science*, 1(3):272–282, 2000.

[68] D. D. Woods, J. Tittle, M. Feil, and A. Roesler. Envisioning human-robot coordination in future operations. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 34(6):749–756, November 2004.